

34/PRTS

- 1 -

10/537798

JC17 Rec'd PCT/PTO 07 JUN 2005

**Proteins involved in the regulation of energy homeostasis****Description**

This invention relates to the use of *four wheel drive* (Gadfly Accession Number CG7004; referred to as *fwd*), *Protein phosphatase 2C* (Gadfly Accession Number CG2984; referred to as *Pp2C1*), *Adenylate kinase 3* (Gadfly Accession Number CG6612; referred to as *Adk3*), Gadfly Accession Number CG3860, *Cyclin-dependent kinase 4* (Gadfly Accession Number CG5072; referred to as *Cdk4*), Gadfly Accession Number CG7134, or *Ecdysone-induced protein 75B* (Gadfly Accession Number CG8127; referred to as *Eip75B*) homologous proteins, to the use of nucleic acid sequences encoding these, and to the use of modulators/effectors of the proteins or nucleotides in the diagnosis, study, prevention, and treatment of obesity and/or diabetes and/or metabolic syndrome.

There are several metabolic diseases of human and animal metabolism, e.g., obesity and severe weight loss, that relate to energy imbalance where caloric intake versus energy expenditure is imbalanced. Obesity is one of the most prevalent metabolic disorders in the world. It is still a poorly understood human disease that becomes as a major health problem more and more relevant for western society. Obesity is defined as a body weight more than 20% in excess of the ideal body weight, frequently resulting in a significant impairment of health. It is associated with an increased mortality rate. Besides severe risks of illness, individuals suffering from obesity are often isolated socially.

Obesity is influenced by genetic, metabolic, biochemical, psychological, and behavioral factors, and can be caused by different reasons such as non-insulin dependent diabetes, increase in triglycerides, increase in carbohydrate bound energy and low energy expenditure. As such, it is a complex disorder that must be addressed on several fronts to achieve lasting positive clinical outcome.

- 2 -

Since obesity is not to be considered as a single disorder but as a heterogeneous group of conditions with (potential) multiple causes, it is also characterized by elevated fasting plasma insulin and an exaggerated insulin response to oral glucose intake (Koltermann O.G. et al., (1980) J. Clin. Invest 65: 1272-1284). A clear involvement of obesity in type 2 diabetes mellitus can be confirmed (Kopelman P.G., (2000) Nature 404: 635-643).

Triglycerides and glycogen are used as the body's fuel energy storage. Glycogen is a large branched polymer of glucose residues that is mainly stored in liver and muscle cells. Glycogen synthesis and degradation is central to the control of the blood glucose level.

Triglycerides are stored in the cytoplasm of adipocytes. Adipocytes are specialized for the synthesis, storage and mobilization of triglycerides. The glycogen and triglyceride metabolism is highly regulated and their interplay is essential for the energy homeostasis of the body. A high glucose level in the adipose cell results in the synthesis of triglycerides as fuel store. A low intracellular glucose level leads to a release of fatty acids, which can be used as substrates for the beta-oxidation to generate energy. Glycogen levels in cells are more variable than triglyceride levels because the turnover of glycogen is higher. Triglycerides are used as long term energy donors once the glycogen stores run low.

Insulin amongst other hormones plays a key role in the regulation of the fuel metabolism. High blood glucose levels stimulate the secretion of insulin by pancreatic beta-cells. Insulin leads to the storage of glycogen and triglycerides and to the synthesis of proteins. The entry of glucose into muscles and adipose cells is stimulated by insulin.

In patients who suffer from diabetes mellitus either the amount of insulin produced by the pancreatic islet cells is too low (Diabetes Type 1 or insulin dependent diabetes mellitus IDDM) or liver and muscle cells lose their ability to respond to normal blood insulin levels (insulin resistance). In the next stage

- 3 -

pancreatic cells become unable to produce sufficient amounts of insulin (Diabetes Type II or non insulin dependent diabetes mellitus NIDDM).

5 Hyperlipidemia and elevation of free fatty acids correlate clearly with the metabolic syndrome, which is defined as the linkage between several diseases, including obesity and insulin resistance. This often occurs in the same patients and are major risk factors for development of type 2 diabetes and cardiovascular disease. It was suggested that the control of lipid levels and glucose levels is required to treat type 2 diabetes, heart disease, and  
10 other occurrences of metabolic syndrome (see, for example, Santomauro A.T. et al., (1999) Diabetes, 48: 1836-1841 and Lakka H.M. et al., (2002) JAMA 288: 2709-2716).

Pancreatic beta-cells secrete insulin in response to blood glucose levels.  
15 Insulin amongst other hormones plays a key role in the regulation of the fuel metabolism. Insulin leads to the storage of glycogen and triglycerides and to the synthesis of proteins. The entry of glucose into muscles and adipose cells is stimulated by insulin. In patients who suffer from diabetes mellitus type I or LADA (latent autoimmune diabetes in adults (Pozzilli & Di Mario, 2001, Diabetes Care. 8: 1460-1467) beta-cells are being destroyed due to  
20 autoimmune attack. The amount of insulin produced by the remaining pancreatic islet cells is too low, resulting in elevated blood glucose levels (hyperglycemia). In diabetes type 2 liver and muscle cells lose their ability to respond to normal blood insulin levels (insulin resistance). High blood  
25 glucose levels (and also high blood lipid levels) in turn lead to an impairment of beta-cell function and to an increase in beta-cell apoptosis.

Diabetes is a very disabling disease, because today's common anti-diabetic  
30 drugs do not control blood sugar levels well enough to completely prevent the occurrence of high and low blood sugar levels. Out of range blood sugar levels are toxic and cause long-term complications like for example nephropathy, retinopathy, neuropathy and peripheral vascular disease. There is also a host of related conditions, such as obesity, hypertension, heart

- 4 -

disease and hyperlipidemia , for which persons with diabetes are substantially at risk.

5 Apart from the impaired quality of life for the patients, the treatment of diabetes and its long term complications presents an enormous financial burden to our healthcare systems with rising tendency. Thus, for the treatment of, type I and type II diabetes as well as for latent autoimmune diabetes in adults (LADA) there is a strong need in the art to identify factors that induce regeneration of pancreatic insulin producing beta-cells. These  
10 factors could restore normal function of the endocrine pancreas once its function is impaired or event could prevent the development or progression of diabetes type I, diabetes type II, or LADA.

The concept of 'metabolic syndrome' (syndrome x, insulin-resistance syndrome, deadly quartet) was first described 1966 by Camus and reintroduced  
15 1988 by Reaven (Camus J.P., (1966) Rev Rhum Mal Osteoartic 33: 10-14; Reaven G.M. et al., (1988) Diabetes, 37: 1595-1607). Today metabolic syndrome is commonly defined as clustering of cardiovascular risk factors like hypertension, abdominal obesity, high blood levels of triglycerides and fasting glucose as well as low blood levels of HDL cholesterol. Insulin resistance  
20 greatly increases the risk of developing the metabolic syndrome (Reaven G., (2002) Circulation 106: 286-288). The metabolic syndrome often precedes the development of type II diabetes and cardiovascular disease (Lakka H.M. et al., 2002, supra).

25

Hyperlipidemia and elevation of free fatty acids correlate clearly with the metabolic syndrome, which is defined as the linkage between several diseases, including obesity an insulin resistance. This often occurs in the same patients and is a major risk factor for development of Type 2 diabetes and  
30 cardiovascular disease. It was suggested that the control of lipid levels and glucose levels is required to treat Type 2 Diabetes, heart disease, and other occurances of metabolic syndrome (see, for example, Santomauro A.T. et al., (1999) Diabetes, 48: 1836-1841).

- 5 -

The molecular factors regulating food intake and body weight balance are incompletely understood. Even if several candidate genes have been described which are supposed to influence the homeostatic system(s) that regulate body mass/weight, like leptin or the peroxisome proliferator-activated receptor-gamma co-activator, the distinct molecular mechanisms and/or molecules influencing obesity or body weight/body mass regulations are not known. In addition, several single-gene mutations resulting in obesity have been described in mice, implicating genetic factors in the etiology of obesity (Friedman J.M. and Leibel R.L., (1990) Cell 69: 217-220). In the ob mouse a single gene mutation (obese) results in profound obesity, which is accompanied by diabetes (Friedman J.M. et. al., (1991) Genomics 11: 1054-1062).

Therefore, the technical problem underlying the present invention was to provide for means and methods for modulating (pathological) metabolic conditions influencing body-weight regulation and/or energy homeostatic circuits. The solution to said technical problem is achieved by providing the embodiments characterized in the claims.

Accordingly, the present invention relates to novel functions of proteins and nucleic acids encoding these in body-weight regulation, energy homeostasis, metabolism, and obesity. Further new compositions are provided that are useful in diagnosis, treatment, and prognosis of metabolic diseases and disorders as described.

So far, it has not been described that a protein of the invention or a homologous protein is involved in the regulation of energy homeostasis and body-weight regulation and related disorders, and thus, no functions in metabolic diseases and dysfunctions and other diseases as listed above have been discussed.

Before the present proteins, nucleotide sequences, and methods are

- 6 -

described, it is understood that this invention is not limited to the particular methodology, protocols, cell lines, vectors, and reagents described as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit  
5 the scope of the present invention that will be limited only by the appended claims. Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or  
10 testing of the present invention, the preferred methods, devices, and materials are now described. All publications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the cell lines, vectors, and methodologies that are reported in the publications which might be used in connection with the invention. Nothing herein is to be construed as  
15 an admission that the invention is not entitled to antedate such disclosure.

The present invention discloses that *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous proteins (herein referred to as "proteins of the invention" or "a protein of the invention") are regulating the energy  
20 homeostasis and fat metabolism, especially the metabolism and storage of triglycerides, and polynucleotides, which identify and encode the proteins disclosed in this invention. The invention also relates to vectors, host cells, and recombinant methods for producing the polypeptides and polynucleotides of the invention. The invention also relates to the use of these compounds and  
25 effectors/modulators thereof, e.g. antibodies, biologically active nucleic acids, such as antisense molecules, RNAi molecules or ribozymes, aptamers, peptides or low-molecular weight organic compounds recognizing said polynucleotides or polypeptides, in the diagnosis, study, prevention, and treatment of metabolic diseases or dysfunctions, including obesity, diabetes  
30 mellitus and/or metabolic syndrome, as well as related disorders such as eating disorder, cachexia, pancreatitis, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, or liver fibrosis.

- 7 -

The synthesis of phosphatidylinositol 4,5-bisphosphate [PtdIns(4,5)P<sub>2</sub>], the immediate precursor of intracellular signals generated by calcium-mobilizing hormones and growth factors, is initiated by the conversion of phosphatidylinositol to phosphatidylinositol 4-phosphate [PtdIns(4)P] by phosphatidylinositol 4-kinase (PtdIns 4-kinase) (Meyers R. and Cantley L.C., (1997) J. Biol. Chem. 272: 4384-4390). Targeting of Golgi-specific pleckstrin homology domains of oxysterol binding protein (OSBP) involves both PtdIns 4-kinase-dependent and -independent components (Levine T.P. and Munro S., (2002) Curr Biol 12: 695-704).

Biochemical analyses indicated that phosphatidyl inositol 4-kinase beta (PI4Kbeta, PIK4CB) is a type III enzyme that is sensitive to wortmannin (Meyers R. and Cantley L.C., supra). PI4Kbeta is localized in the cytosol and also present in the Golgi region (Wong K. et al., (1997) J Biol Chem 272: 13236-13241). PI4Kbeta is ubiquitously expressed, with highest expression in heart, pancreas, and skeletal muscle.

PI4Kbeta (PIK4CB, PI4K type III) was primarily found in the Golgi, but it was also present in the walls of numerous large perinuclear vesicles. Co-expression of a catalytically inactive PI4Kbeta inhibited the development of this vesicular phenotype. PI4Kbeta is involved in vesicular trafficking (Zhao X. et al., (2001) J Biol Chem 276: 40183-40189).

As shown in this invention magnesium-dependent protein phosphatase 1 D (PPM1D) is the human homolog of Drosophila protein phosphatase 2C (Pp2C1). The human wildtype p53-induced phosphatase 1 (Wip1; GenBank symbol PPM1D) gene encodes a type 2C protein phosphatase (PP2C) that is induced by ionizing radiation in a p53-dependent manner. The murine Wip1 mRNA is expressed ubiquitously in adult and embryonic tissues, though expression in the testis was much higher than in other tissues (Choi J. et al., (2000) Genomics 64: 298-306). Mice deficient for the wild-type p53-induced phosphatase gene (Wip1) exhibit defects in

- 8 -

reproductive organs, immune function, and cell cycle control (Choi J. et al., (2002) Mol Cell Biol, 22: 1094-1105). Inactivation of the p38 MAPK through PPM1D overexpression resulting from PPM1D amplification contributes to the development of human cancers by suppressing p53 activation (Bulavin D.V. et al, (2002) Nat Genet 31: 210-215).

Adenylate kinases regulate the adenine and guanine nucleotide compositions within a cell by catalyzing the reversible transfer of phosphate group among these nucleotides. Three isozymes of adenylate kinase have been identified in vertebrates, adenylate isozyme 1 (AK1), 2 (AK2) and 3 (AK3). Expression of these isozymes is tissue-specific and developmentally regulated.

AK3 is present in the mitochondrial matrix and prefers GTP over ATP as the substrate (Wilson D.E. et al. (1976) Ann Hum Genet 39: 305-313). In the failing myocardium, phosphotransfer activities of creatine kinase, adenylate kinase, 3-phosphoglycerate kinase and pyruvate kinase, which collectively deliver ATP and remove ADP from myofibrillar ATPases, were depressed, when compared to normal controls. As these enzymatic systems are collectively required for adequate delivery of high-energy phosphoryl to, and removal of end-products from, cellular ATPases, the cumulative deficit in their flux capacities may provide a bioenergetic basis for impaired contraction-relaxation in the failing heart (Dzeja P.P. et al., (1999) Mol Cell Biochem 201: 33-40).

The OSBPL1A and OSBP2 genes encode members of the oxysterol-binding protein (OSBP) family, a group of intracellular lipid receptors. Most members contain an N-terminal pleckstrin homology domain and a highly conserved C-terminal OSBP-like sterol-binding domain, although some members contain only the sterol-binding domain (e.g. OSBP2). The OSBP2 protein contains only the sterol-binding domain. In vitro studies have shown that the encoded protein can bind strongly to phosphatic acid and weakly to phosphatidylinositol 3-phosphate, but cannot bind to 25-hydroxycholesterol. The protein associates with the Golgi apparatus.



ORP2, an oxysterol binding protein related protein, is a regulator of cellular sterol homeostasis and intracellular membrane trafficking (Laitinen S. et al. (2002) J Lipid Res 43: 245-255 ).

5

Cyclin-dependent kinase 6 (CDK6) and cyclin-dependent kinase 4 (CDK4) are members of the cyclin-dependent protein kinase (CDK) family. CDK family members are highly similar to the gene products of *Saccharomyces cerevisiae* cdc28, and *Schizosaccharomyces pombe* cdc2, and are known to be important regulators of cell cycle progression. CDK6 and CDK4 are catalytic subunits of the protein kinase complex that is important for cell cycle G1 phase progression and G1/S transition, these kinases have been shown to phosphorylate, and thus regulate the activity of, tumor suppressor protein retinoblastoma protein (Rb).

15

CDK4(-/-) mice survived embryogenesis and showed growth retardation and reproductive dysfunction associated with hypoplastic seminiferous tubules in the testis and perturbed corpus luteum formation in the ovary. A majority of CDK4(-/-) mice developed diabetes mellitus by 6 weeks, associated with degeneration of pancreatic islets. (Tsutsui T. et al., (1999) Mol Cell Biol 19: 7011-7019).

20

The protein encoded by CDC14 is a member of the dual specificity protein tyrosine phosphatase family. This protein is highly similar to *Saccharomyces cerevisiae* Cdc14, a protein tyrosine phosphatase involved in the exit of cell mitosis and initiation of DNA replication, which suggests the role in cell cycle control. This protein has been shown to interact with and dephosphorylates tumor suppressor protein p53, and is thought to regulate the function of p53. Alternative splice of this gene results in 3 transcript variants encoding distinct isoforms.

30

CDC14A gene is expressed as 1.8- and 4.4-kb mRNAs in all tissues, with the strongest expression in kidney, heart, and skeletal muscle, CDC14B

- 10 -

shows ubiquitous expression. In budding yeast, the Cdc14p phosphatase activates mitotic exit by dephosphorylation of specific cyclin-dependent kinase (Cdk) substrates and seems to be regulated by sequestration in the nucleolus until its release in mitosis. Herein, we have analyzed the two human homologs of Cdc14p, hCdc14A and hCdc14B. We demonstrate that the human Cdc14A phosphatase is selective for Cdk substrates in vitro and that although the protein abundance and intrinsic phosphatase activity of hCdc14A and B vary modestly during the cell cycle, their localization is cell cycle regulated. hCdc14A dynamically localizes to interphase but not mitotic centrosomes, and hCdc14B localizes to the interphase nucleolus. These distinct patterns of localization suggest that each isoform of human Cdc14 likely regulates separate cell cycle events. In addition, hCdc14A overexpression induces the loss of the pericentriolar markers pericentrin and gamma-tubulin from centrosomes. Overproduction of hCdc14A also causes mitotic spindle and chromosome segregation defects, defective karyokinesis, and a failure to complete cytokinesis. Thus, the hCdc14A phosphatase appears to play a role in the regulation of the centrosome cycle, mitosis, and cytokinesis, thereby influencing chromosome partitioning and genomic stability in human cells (Kaiser B.K. et al. (2002) Mol Biol Cell 13: 2289-2300).

Cdc14A phosphatase interacts with interphase centrosomes, and that this interaction is independent of microtubules and Cdc14A phosphatase activity, but requires active nuclear export. Disrupting the nuclear export signal (NES) led to Cdc14A being localized in nucleoli, which in unperturbed cells selectively contain Cdc14B. Conditional overproduction of Cdc14A, but not its phosphatase-dead or NES-deficient mutants, or Cdc14B, resulted in premature centrosome splitting and formation of supernumerary mitotic spindles. In contrast, downregulation of endogenous Cdc14A by short inhibitory RNA duplexes (siRNA) induced mitotic defects including impaired centrosome separation and failure to undergo productive cytokinesis. Consequently, both overexpression and downregulation of Cdc14A caused aberrant chromosome partitioning into daughter cells. These results indicate

- 11 -

thatCdc14A is a physiological regulator of the centrosome duplication cycle, which, when disrupted, can lead to genomic instability in mammalian cells (Mailand N. et al., (2002) Nat Cell Biol 4: 317-322). CDC14 is involved in the control of the cell cycle (review Oliferenko S. and Balasubramanian M.K. (2001) Curr Biol 11: R872-874). Cdc14 activates protein kinase cdc15 to promote mitotic exit in budding yeast (Jaspersen S.L. and Morgan D.O., (2000) Curr Biol 10: 615-618).

Nuclear receptor subfamily 1, group D, member 1 (NR1D1) and member 2 (NR1D2) are members of the nuclear receptor superfamily of ligand-activated transcription factors. NR1D1 is a thyroid/steroid hormone receptor and functions as a transcriptional regulator. The NR1D1 gene encodes the orphan receptor Rev-ErbA alpha, and NR1D2 encodes Rev-ErbA beta (Koh Y.S. and Moore D.D., (1999) Genomics 57: 289-292). Rev-ErbA alpha (Rev-Erb) mRNA levels increased during the differentiation of 3T3-L1 cells into adipocytes. Rev-Erb was similarly induced in the related 3T3-F442A cell line but not in nondifferentiating 3T3-C2 cells. Treatment of preadipocytes with retinoic acid inhibited adipocyte differentiation and also prevented Rev-Erb induction (Chawla A. and Lazar M.A., (1993) J Biol Chem 268: 16265-16269). Circadian rhythms are generated by a transcription/translation feedback loop consisting of two limbs, one positive and one negative. The nuclear orphan receptor, Rev-Erb alpha, was identified as a molecular link coupling these two limbs (Alvarez J.D. and Sehgal A., (2002) Dev Cell 3: 150-152). Rev-ErbA beta shows a high expression in particular in the cerebellum, the dentate gyrus of the hippocampus and pituitary gland of adult rats (Enmark E. et al., (1994) Biochem Biophys Res Commun 204: 49-56).

*Fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids encoding the human *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologs as described in Table 1.

- 12 -

The invention particularly relates to a nucleic acid molecule encoding a polypeptide contributing to regulating the energy homeostasis and the metabolism of triglycerides and glycogen, wherein said nucleic acid molecule comprises

- (a) the nucleotide sequence of *Drosophila fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B*, human *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous nucleic acids, particularly the nucleic acids as described in Table 1, and/or a sequence complementary thereto,
- (b) a nucleotide sequence which hybridizes at 50°C in a solution containing 1 x SSC and 0.1% SDS to a sequence of (a),
- (c) a sequence corresponding to the sequences of (a) or (b) within the degeneration of the genetic code,
- (d) a sequence which encodes a polypeptide which is at least 85%, preferably at least 90%, more preferably at least 95%, more preferably at least 98% and up to 99,6% identical to the amino acid sequences of the *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous protein, preferably of the human *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous protein, particularly a protein as described in Table 1,
- (e) a sequence which differs from the nucleic acid molecule of (a) to (d) by mutation and wherein said mutation causes an alteration, deletion, duplication and/or premature stop in the encoded polypeptide or
- (f) a partial sequence of any of the nucleotide sequences of (a) to (e) having a length of 15-25 bases, preferably 25-35 bases, more preferably 35-50 bases and most preferably at least 50 bases.

The invention is based on the finding that *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous proteins and the polynucleotides encoding therefore, are involved in the regulation of triglyceride storage and therefore energy homeostasis. The invention describes the use of compositions comprising *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B*

- 13 -

homologous polypeptides and polynucleotides as well as modulators/effectors thereof for the diagnosis, study, prevention, or treatment of metabolic diseases or dysfunctions, including obesity, diabetes mellitus and/or metabolic syndrome, as well as related disorders such as eating disorder, cachexia, pancreatitis, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones or liver fibrosis.

Accordingly, the present invention relates to genes with novel functions in body-weight regulation, energy homeostasis, metabolism, and obesity, fragments of said genes, polypeptides encoded by said genes or fragments thereof, and effectors e.g. antibodies, biologically active nucleic acids, such as antisense molecules, RNAi molecules or ribozymes, aptamers, peptides or low-molecular weight organic compounds recognizing said polynucleotides or polypeptides.

The ability to manipulate and screen the genomes of model organisms such as the fly *Drosophila melanogaster* provides a powerful tool to analyze biological and biochemical processes that have direct relevance to more complex vertebrate organisms due to significant evolutionary conservation of genes, cellular processes, and pathways (see, for example, Adams M.D. et al., (2000) *Science* 287: 2185-2195). Identification of novel gene functions in model organisms can directly contribute to the elucidation of correlative pathways in mammals (humans) and of methods of modulating them. A correlation between a pathology model (such as changes in triglyceride levels as indication for metabolic syndrome including obesity) and the modified expression of a fly gene can identify the association of the human ortholog with the particular human disease.

In one embodiment, a forward genetic screen is performed in fly displaying a mutant phenotype due to misexpression of a known gene (see, St. Johnston D., (2002) *Nat Rev Genet* 3: 176-188 ; Rorth P., (1996) *Proc Natl Acad Sci USA* 93: 12418-12422). In this invention, we have used a genetic screen to identify mutations of *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B*

- 14 -

genes that cause changes in the body weight which are reflected by a significant change of triglyceride levels. Additionally glycogen levels are analysed.

5 Obese people mainly show a significant increase in the content of triglycerides. Triglycerides are the most efficient storage for energy in cells. In order to isolate genes with a function in energy homeostasis, several thousand proprietary and publicly available EP-lines were tested for their triglyceride content after a prolonged feeding period (see Examples and  
10 Figures for more detail). Lines with significantly changed triglyceride content were selected as positive candidates for further analysis. The change of triglyceride content due to the loss of a gene function suggests gene activities in energy homeostasis in a dose dependent manner that controls the amount of energy stored as triglycerides.

15

The content of triglycerides of a pool of flies with the same genotype was analyzed after prolonged feeding using a triglyceride assay. Male flies homozygous, hemizygous, or heterozygous for the integration of vectors for *Drosophila* EP-lines were analyzed in an assay measuring the triglyceride  
20 contents of these flies, illustrated in more detail in the Examples section. The results of the triglyceride content analysis are shown in Figures 1, 5, 8, 12, 16, 18, and 20, respectively.

Genomic DNA sequences were isolated that are localized adjacent to the EP  
25 vector integration. Using those isolated genomic sequences public databases like Berkeley *Drosophila* Genome Project (GadFly; see also FlyBase (1999) Nucleic Acids Research 27: 85-88) were screened thereby identifying the integration sites of the vectors, and the corresponding genes, described in more detail in the Examples section. The molecular organization of the genes  
30 is shown in Figures 2, 6, 9, 13, 17, 19, and 21, respectively.

The function of *Eip75B* in metabolic disorders is further validated by data obtained from an additional screen. For example, an additional screen using

- 15 -

Drosophila mutants with modifications of the eye phenotype identified a modification of UCP activity by *Eip75B*, thereby leading to an altered mitochondrial activity. These findings suggest the presence of similar activities of these described homologous proteins in humans that provides insight into diagnosis, treatment, and prognosis of metabolic disorders.

The Drosophila genes and proteins encoded thereby with functions in the regulation of triglyceride metabolism were further analysed in publicly available sequence databases (see Examples for more detail) and mammalian homologs were identified.

The function of the mammalian homologs in energy homeostasis was further validated in this invention by analyzing the expression of the transcripts in different tissues and by analyzing the role in adipocyte differentiation. Expression profiling studies (see Examples for more detail) confirm the particular relevance of the proteins of the invention as regulators of energy metabolism in mammals. Further, we show that the proteins of the invention are regulated by fasting and by genetically induced obesity. In this invention, we used mouse models of insulin resistance and/or diabetes, such as mice carrying gene knockouts in the leptin pathway (for example, ob (leptin) or db (leptin receptor) mice) to study the expression of the proteins of the invention. Such mice develop typical symptoms of diabetes, show hepatic lipid accumulation and frequently have increased plasma lipid levels (see Bruning J.C. et al., (1998) Mol. Cell. 2: 559-569).

Microarrays are analytical tools routinely used in bioanalysis. A microarray has molecules distributed over, and stably associated with, the surface of a solid support. The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, antibodies, or other chemical compounds on a substrate. Microarrays of polypeptides, polynucleotides, and/or antibodies have been developed and find use in a variety of applications, such as monitoring gene expression, drug discovery, gene sequencing, gene mapping, bacterial identification, and combinatorial chemistry. One area in particular in

- 16 -

which microarrays find use is in gene expression analysis (see Example 6). Array technology can be used to explore the expression of a single polymorphic gene or the expression profile of a large number of related or unrelated genes. When the expression of a single gene is examined, arrays  
5 are employed to detect the expression of a specific gene or its variants. When an expression profile is examined, arrays provide a platform for identifying genes that are tissue specific, are affected by a substance being tested in a toxicology assay, are part of a signaling cascade, carry out housekeeping functions, or are specifically related to a particular genetic predisposition,  
10 condition, disease, or disorder.

Microarrays may be prepared, used, and analyzed using methods known in the art (see for example, Brennan T.M., (1995) U.S. Patent No. US5474796; Schena M. et al., (1996) Proc. Natl. Acad. Sci. USA 93: 10614-10619;  
15 Baldeschwieler et al., (1995) PCT application WO9525116; Shalon T.D. and Brown P.O., (1995) PCT application WO9535505; Heller R.A. et al., (1997) Proc. Natl. Acad. Sci. USA 94: 2150-2155; Heller M.J. and Tu E., (1997) U.S. Patent No. US5605662). Various types of microarrays are well known and thoroughly described in Schena M., ed. (1999); DNA Microarrays: A Practical  
20 Approach, Oxford University Press, London.

Oligonucleotides or longer fragments derived from any of the polynucleotides described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques, which monitor the relative  
25 expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene  
30 expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic



- 17 -

agents, which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

As determined by microarray analysis, phosphatidylinositol 4-kinase, catalytic, beta polypeptide (PIK4CB), adenylate kinase 3 like 1 (AK3L1), adenylate kinase 3 (AK3), oxysterol binding protein-like 1A (OSBPL1A), and oxysterol binding protein-like 2 (OSBPL2) are strong candidates for the manufacture of a pharmaceutical composition and a medicament for the treatment of conditions related to human metabolism, such as obesity, diabetes, and/or metabolic syndrome.

In another embodiment, we found that triglyceride levels were significantly increased in Nr1d-1 overexpressing cells during adipocyte differentiation. Glycogen levels in cells were more variable than triglyceride levels because the turnover of glycogen is higher. Glucose is taken up by the cells rapidly and stored in the form of glycogen. This energy storage is then used as a first quick response to the metabolic demands of the cell. During differentiation of adipocytes, glucose uptake in Nr1d-1 overexpressing cells is significantly increased. In another embodiment, triglyceride levels were significantly decreased in cells without functional Nr1d1 during adipocyte differentiation. Insulin stimulated lipid synthesis levels were significantly decreased during adipocyte differentiation in cells without functional Nr1d1. These findings show that the Nr1d1 plays a key role in central metabolic pathways in the cell.

The invention also encompasses polynucleotides that encode the proteins of the invention and homologous proteins. Accordingly, any nucleic acid sequence, which encodes the amino acid sequences of the proteins of the invention and homologous proteins, can be used to generate recombinant molecules that express the proteins of the invention and homologous proteins. In a particular embodiment, the invention encompasses a nucleic acid encoding *Drosophila fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* or human *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous proteins, preferably a human homologous protein as described in

- 18 -

Table 1; referred to herein as the proteins of the invention. It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding the proteins, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. The invention contemplates each and every possible variation of nucleotide sequence that can be made by selecting combinations based on possible codon choices.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those of the polynucleotide encoding the proteins of the invention, under various conditions of stringency. Hybridization conditions are based on the melting temperature ( $T_m$ ) of the nucleic acid binding complex or probe, as described in Wahl G.M. et al. (1987: Methods Enzymol. 152: 399-407) and Kimmel A.R. (1987; Methods Enzymol. 152: 507-511), and may be used at a defined stringency. Preferably, hybridization under stringent conditions means that after washing for 1 h with 1 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at 65°C, particularly for 1 h in 0.2 x SSC and 0.1% SDS at 50°C, preferably at 55°C, more preferably at 62°C and most preferably at 65°C, a positive hybridization signal is observed. Altered nucleic acid sequences encoding the proteins which are encompassed by the invention include deletions, insertions or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent protein.

The encoded proteins may also contain deletions, insertions or substitutions of amino acid residues, which produce a silent change and result in functionally equivalent proteins. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of the protein is retained. Furthermore, the invention relates to peptide fragments of the proteins or derivatives thereof such as cyclic peptides, retro-inverso peptides or peptide mimetics having a length of at least 4,

- 19 -

preferably at least 6 and up to 50 amino acids.

Also included within the scope of the present invention are alleles of the genes encoding the proteins of the invention. As used herein, an 'allele' or 'allelic sequence' is an alternative form of the gene, which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structures or function may or may not be altered. Any given gene may have none, one or many allelic forms. Common mutational changes, which give rise to alleles, are generally ascribed to natural deletions, additions or substitutions of nucleotides. Each of these types of changes may occur alone or in combination with the others, one or more times in a given sequence.

The nucleic acid sequences encoding the proteins of the invention and homologous proteins may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed, 'restriction-site' PCR, uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar G. et al., (1993) PCR Methods Applic. 2: 318-322). Inverse PCR may also be used to amplify or extend sequences using divergent primers based on a known region (Triglia T. et al., (1988) Nucleic Acids Res. 16: 8186). Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom M. et al., (1991) PCR Methods Applic. 1: 111-119). Another method which may be used to retrieve unknown sequences is that of Parker J.D. et al., (1991) Nucleic Acids Res. 19: 3055-3060. Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries to walk in genomic DNA (Clontech, Palo Alto, Calif.). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

In order to express a biologically active protein, the nucleotide sequences

- 20 -

encoding the proteins or functional equivalents, may be inserted into appropriate expression vectors, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods, which are well known to those skilled in the art, may be used to construct expression vectors containing sequences encoding the proteins and the appropriate transcriptional and translational control elements. Regulatory elements include for example a promoter, an initiation codon, a stop codon, a mRNA stability regulatory element, and a polyadenylation signal. Expression of a polynucleotide can be assured by (i) constitutive promoters such as the cytomegalovirus (CMV) promoter/enhancer region, (ii) tissue specific promoters such as the insulin promoter (see, Soria B. et al., (2000) Diabetes 49: 157-162), SOX2 gene promoter (see Li M. et al., (1998) Curr. Biol. 8: 971-974), Msi-1 promoter (see Sakakibara S. and Okano H., (1997) J. Neuroscience 17: 8300-8312), alpha-cardia myosin heavy chain promoter or human atrial natriuretic factor promoter (Klug M.G. et al., (1996) J. Clin. Invest 98: 216-224; Wu J. et al., (1989) J. Biol. Chem. 264: 6472-6479) or (iii) inducible promoters such as the tetracycline inducible system. Expression vectors can also contain a selection agent or marker gene that confers antibiotic resistance such as the neomycin, hygromycin or puromycin resistance genes. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al., (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y. and Ausubel, F.M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

In a further embodiment of the invention, natural, modified or recombinant nucleic acid sequences encoding the proteins of the invention and homologous proteins may be ligated to a heterologous sequence to encode a fusion protein.

A variety of expression vector/host systems may be utilized to contain and express sequences encoding the proteins or fusion proteins. These include,

- 21 -

but are not limited to, micro-organisms such as bacteria transformed with recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus, adenovirus, adeno-associated virus, lentivirus, retrovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or PBR322 plasmids); or animal cell systems.

10 The presence of polynucleotide sequences of the invention in a sample can be detected by DNA-DNA or DNA-RNA hybridization and/or amplification using probes or portions or fragments of said polynucleotides. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences specific for the gene to detect transformants containing DNA or RNA encoding the corresponding protein. As used herein 'oligonucleotides' or 'oligomers' refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplifier.

20 A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting polynucleotide sequences include oligo-labeling, nick translation, end-labeling of RNA probes, PCR amplification using a nucleotide, or enzymatic synthesis. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, Mich.); Promega (Madison Wis.); and U.S. Biochemical Corp., (Cleveland, Ohio).

30 The presence of proteins of the invention in a sample can be determined by immunological methods or activity measurement. A variety of protocols for detecting and measuring the expression of proteins, using either polyclonal or monoclonal antibodies specific for the protein or reagents for determining

- 22 -

protein activity are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on the protein is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, Minn.) and Maddox, D.E. et al. (1983; J. Exp. Med. 158: 1211-1226).

Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent or chromogenic agents as well as substrates, co-factors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding a protein of the invention may be cultured under conditions suitable for the expression and recovery of said protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence or/and the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides, which encode the protein may be designed to contain signal sequences, which direct secretion of the protein through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding the protein to nucleotide sequence encoding a polypeptide domain, which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAG extension/affinity purification system (Immunex Corp., Seattle, Wash.) The inclusion of cleavable linker sequences such as those specific for Factor XA or Enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and the desired protein may be used to facilitate purification.

## Diagnostics and Therapeutics

The data disclosed in this invention show that the nucleic acids and proteins of the invention and modulator/effector molecules thereof are useful in diagnostic and therapeutic applications implicated, for example, but not limited to, metabolic syndrome including obesity, diabetes mellitus, eating disorder, cachexia, pancreatitis, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, gallstones, or liver fibrosis. Hence, diagnostic and therapeutic uses for the proteins and nucleic acids of the invention and modulators/effectors thereof are, for example but not limited to, the following: (i) protein therapy, (ii) small molecule drug target, (iii) antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) diagnostic and/or prognostic marker, (v) gene therapy (gene delivery/gene ablation), (vi) research tools, and (vii) tissue regeneration in vitro and in vivo (regeneration for all these tissues and cell types composing these tissues and cell types derived from these tissues).

The nucleic acids and proteins of the invention and modulators/effectors thereof are useful in diagnostic and therapeutic applications implicated in various applications as described below. For example, but not limited to, cDNAs encoding the proteins of the invention and particularly their human homologues may be useful in gene therapy, and the proteins of the invention and particularly their human homologues may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the present invention will have efficacy for treatment of patients suffering from, for example, but not limited to, in metabolic disorders as described above.

The nucleic acids of the invention or fragments thereof, may further be useful in diagnostic applications, wherein the presence or amount of the nucleic acids or the proteins are to be assessed. Further antibodies that bind immunospecifically to the novel substances of the invention may be used in therapeutic or diagnostic methods.

- 24 -

For example, in one aspect, antibodies, which are specific for the proteins of the invention and homologous proteins, may be used directly as an effector, e.g. an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express the protein.

5 The antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralising antibodies, (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

10 For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with the protein or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. It is preferred that the peptides, fragments or  
15 oligopeptides used to induce antibodies to the protein have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids.

20 Monoclonal antibodies to the proteins may be prepared using any technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Köhler, G. and Milstein C., (1975) Nature 256: 495-497; Kozbor, D. et al.,  
25 (1985) J. Immunol. Methods 81: 31-42; Cote, R.J. et al., (1983) Proc. Natl. Acad. Sci. 80: 2026-2030; Cole, S.P. et al., (1984) Mol. Cell Biochem. 62: 109-120).

30 In addition, techniques developed for the production of 'chimeric antibodies', the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S.L. et al., (1984) Proc. Natl. Acad. Sci. USA 81: 6851-6855; Neuberger, M.S. et al., (1984) Nature 312: 604-608; Takeda, S. et al., (1985)



- 25 -

Nature 314: 452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce single chain antibodies specific for the proteins of the invention and homologous proteins. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries (Kang, A.S. et al., (1991) Proc. Natl. Acad. Sci. USA 88: 11120-11123). Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86: 3833-3837; Winter, G. and Milstein C., (1991) Nature 349: 293-299).

Antibody fragments which contain specific binding sites for the proteins may also be generated. For example, such fragments include, but are not limited to, the F(ab')<sub>2</sub> fragments which can be produced by Pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse W.D. et al. (1989) Science 254: 1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding and immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between the protein and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering protein epitopes are preferred, but a competitive binding assay may also be employed (Maddox, supra).

In another embodiment of the invention, the polynucleotides of the invention or fragments thereof or nucleic acid modulator/effector molecules such as

- 26 -

aptamers, antisense molecules, aptamers, RNAi molecules or ribozymes may be used for therapeutic purposes. In one aspect, aptamers, i.e. nucleic acid molecules, which are capable of binding to a protein of the invention and modulating its activity, may be generated by a screening and selection  
5 procedure involving the use of combinatorial nucleic acid libraries.

In a further aspect, antisense molecules may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides  
10 encoding the proteins of the invention and homologous proteins. Thus, antisense molecules may be used to modulate protein activity or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding  
15 the proteins. Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods, which are well known to those skilled in the art, can be used to construct recombinant vectors, which will express antisense molecules  
20 complementary to the polynucleotides of the genes encoding the proteins of the invention and homologous proteins. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra). Genes encoding the proteins of the invention and homologous proteins can be turned off by transforming a cell or tissue with expression vectors, which express high levels  
25 of polynucleotides that encode the proteins of the invention and homologous proteins or fragments thereof. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient  
30 expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by

- 27 -

designing antisense molecules, e.g. DNA, RNA or nucleic acid analogues such as PNA, to the control regions of the genes encoding the proteins of the invention and homologous proteins, i.e., the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g.,  
5 between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it cause inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors or regulatory molecules. Recent therapeutic advances using triplex DNA have  
10 been described in the literature (Gee J.E. et al., (1994) Gene 149: 109-114; Huber B.E. and Carr B.I., Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y.). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

15 Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples, which may be  
20 used, include engineered hammerhead motif ribozyme molecules that can be specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding the proteins of the invention and homologous proteins. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the  
25 following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by  
30 testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Nucleic acid modulator/effector molecules, e.g. antisense molecules and

- 28 -

ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences. Such DNA sequences may be incorporated into a variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells or tissues. RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule or modifications in the nucleobase, sugar and/or phosphate moieties, e.g. the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of non-traditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods, which are well known in the art. Any of the therapeutic methods described above may be applied to any suitable subject including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such

- 29 -

pharmaceutical compositions may consist of the nucleic acids and the proteins of the invention and homologous nucleic acids or proteins, antibodies to the proteins of the invention and homologous proteins, mimetics, agonists, antagonists or inhibitors of the proteins of the invention and homologous proteins or nucleic acids. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone or in combination with other agents, drugs or hormones. The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries, which facilitate processing of the active compounds into preparations, which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art. For any compounds, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of preadipocyte cell lines or in animal models, usually mice, rabbits, dogs or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. A therapeutically effective dose refers to that amount of active

- 30 -

ingredient, for example the nucleic acids or the proteins of the invention or fragments thereof or antibodies, which is sufficient for treating a specific condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between therapeutic and toxic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions, which exhibit large therapeutic indices, are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage from employed, sensitivity of the patient, and the route of administration. The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors, which may be taken into account, include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week or once every two weeks depending on half-life and clearance rate of the particular formulation. Normal dosage amounts may vary from 0.1 to 100,000  $\mu\text{g}$ , up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

In another embodiment, antibodies which specifically bind to the proteins may be used for the diagnosis of conditions or diseases characterized by or

- 31 -

associated with over- or under-expression of the proteins of the invention and homologous proteins or in assays to monitor patients being treated with the proteins of the invention and homologous proteins, or effectors thereof, e.g. agonists, antagonists, or inhibitors. Diagnostic assays include methods which  
5 utilize the antibody and a label to detect the protein in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used several of which are described above.

10 A variety of protocols including ELISA, RIA, and FACS for measuring proteins are known in the art and provide a basis for diagnosing altered or abnormal levels of gene expression. Normal or standard values for gene expression are established by combining body fluids or cell extracts taken from normal  
15 mammalian subjects, preferably human, with antibodies to the protein under conditions suitable for complex formation. The amount of standard complex formation may be quantified by various methods, but preferably by photometric means. Quantities of protein expressed in control and disease, samples e.g. from biopsied tissues are compared with the standard values. Deviation  
20 between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides specific for the proteins of the invention and homologous proteins may be used for diagnostic  
25 purposes. The polynucleotides, which may be used, include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which gene expression may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence,  
30 and excess gene expression, and to monitor regulation of protein levels during therapeutic intervention.

In one aspect, hybridization with probes which are capable of detecting

- 32 -

polynucleotide sequences, including genomic sequences, encoding the proteins of the invention and homologous proteins or closely related molecules, may be used to identify nucleic acid sequences which encode the respective protein. The hybridization probes of the subject invention may be  
5 DNA or RNA and are preferably derived from the nucleotide sequence of the polynucleotide encoding the proteins of the invention or from a genomic sequence including promoter, enhancer elements, and introns of the naturally occurring gene. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as  $^{32}\text{P}$  or  $^{35}\text{S}$  or enzymatic labels, such  
10 as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences specific for the proteins of the invention and homologous nucleic acids may be used for the diagnosis of conditions or  
15 diseases, which are associated with the expression of the proteins. Examples of such conditions or diseases include, but are not limited to, metabolic diseases and disorders, including obesity and diabetes. Polynucleotide sequences specific for the proteins of the invention and homologous proteins may also be used to monitor the progress of patients receiving treatment for  
20 metabolic diseases and disorders, including obesity and diabetes. The polynucleotide sequences may be used qualitative or quantitative assays, e.g. in Southern or Northern analysis, dot blot or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect altered gene  
25 expression.

In a particular aspect, the nucleotide sequences specific for the proteins of the invention and homologous nucleic acids may be useful in assays that detect activation or induction of various metabolic diseases or dysfunctions,  
30 including metabolic syndrome, obesity, or diabetes mellitus as well as related disorders such as eating disorder, cachexia, pancreatitis, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, or liver fibrosis. The nucleotide sequences may be



- 33 -

labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. The presence of altered levels of nucleotide sequences encoding the proteins of the invention and homologous proteins in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disease associated with expression of the proteins of the invention and homologous proteins, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence or a fragment thereof, which is specific for the nucleic acids encoding the proteins of the invention and homologous nucleic acids, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease. Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that, which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to metabolic diseases such as described above the presence of an unusual amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease or may provide a means for detecting the disease prior to the appearance of actual clinical

- 34 -

symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the metabolic diseases and disorders.

5

Additional diagnostic uses for oligonucleotides designed from the sequences encoding the proteins of the invention and homologous proteins may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5'.fwdarw.3') and another with antisense (3'.rarw.5'), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantification of closely related DNA or RNA sequences.

15

In another embodiment of the invention, the nucleic acid sequences may also be used to generate hybridization probes, which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH, FACS or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price C.M., (1993) Blood Rev. 7: 127-134, and Trask B.J., (1991) Trends Genet. 7: 149-154. FISH (as described in Verma R.S. and Babu A., (1989) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, N.Y.). The results may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265: 1981f). Correlation between the location of the gene encoding the proteins of the invention on a physical chromosomal map and a specific disease or predisposition to a specific disease, may help to delimit the region of DNA associated with that genetic disease.

20

25

30

- 35 -

The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals. An analysis of polymorphisms, e.g. single nucleotide polymorphisms may be carried out. Further, in situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti R.A. et al., (1988) Nature 336: 577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequences of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier or affected individuals.

In another embodiment of the invention, the proteins of the invention, their catalytic or immunogenic fragments or oligopeptides thereof, an in vitro model, a genetically altered cell or animal, can be used for screening libraries of compounds, e.g. peptides or low molecular weight organic compounds, in any of a variety of drug screening techniques. One can identify modulators/effectors, e.g. receptors, enzymes, proteins, ligands, or substrates that bind to, modulate or mimic the action of one or more of the proteins of the invention. The protein or fragment thereof employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between the protein of the invention and the agent tested, may be measured. Agents could also, either directly or indirectly, influence the

- 36 -

activity of the proteins of the invention.

For example the phosphatase activity of the proteins of the invention could be measured in vitro by using recombinantly expressed and purified CG7134  
5 or *Pp2C1* homologous phosphatase or fragments thereof by making use of artificial substrates well known in the art, i.e. but not exclusively DiFMUP (Molecular Probes, Eugene, Oregon), which are converted to fluorophores or chromophores upon dephosphorylation. Alternatively, the dephosphorylation of physiological substrates of the phosphatases could be measured by  
10 making use of any of the well known screening technologies suitable for the detection of the phosphorylation status of their physiological substrates. For example, but not exclusively, the phosphorylation status of peptides derived from their physiological substrates can be monitored by binding of phospho-side specific antibodies resulting in an increase of the polarization of the  
15 complex.

In vivo, the enzymatic kinase activity of the unmodified polypeptides of *Cdk4*, *fwd*, or *Adk3* homologous kinase towards a substrate can be enhanced by appropriate stimuli, triggering the phosphorylation of *Cdk4*, *fwd*, or *Adk3*  
20 homologous kinase. This may be induced in the natural context by extracellular or intracellular stimuli, such as signaling molecules or environmental influences. One may generate a system containing activated *Cdk4*, *fwd*, or *Adk3* homologous kinase, may it be an organism, a tissue, a culture of cells or cell-free environment, by exogenously applying this stimulus  
25 or by mimicking this stimulus by a variety of the techniques, some of them described further below. A system containing activated *Cdk4*, *fwd*, or *Adk3* homologous kinase may be produced (i) for the purpose of diagnosis, study, prevention, and treatment of diseases and disorders related to body-weight regulation and thermogenesis, for example, but not limited to, metabolic  
30 diseases or dysfunctions, including metabolic syndrome, obesity, or diabetes, as well as related disorders such as eating disorder, cachexia, hypertension, coronary heart disease, hypercholesterolemia, dyslipidemia, osteoarthritis, gallstones, or liver fibrosis.

- 37 -

In addition activity of *fwd*, *Pp2C1*, *Adk3*, *CG3860*, *Cdk4*, *CG7134*, or *Eip75B* homologous proteins against their physiological substrate(s) or derivatives thereof could be measured in cell-based assays. Agents may also interfere  
5 with posttranslational modifications of the protein, such as phosphorylation and dephosphorylation, farnesylation, palmitoylation, acetylation, alkylation, ubiquitination, proteolytic processing, subcellular localization and degradation. Moreover, agents could influence the dimerization or oligomerization of the proteins of the invention or, in a heterologous manner,  
10 of the proteins of the invention with other proteins, for example, but not exclusively, docking proteins, enzymes, receptors, ion channels, uncoupling proteins, or translation factors. Agents could also act on the physical interaction of the proteins of this invention with other proteins, which are required for protein function, for example, but not exclusively, their  
15 downstream signaling.

Methods for determining protein-protein interaction are well known in the art. For example binding of a fluorescently labeled peptide derived from the interacting protein to the protein of the invention, or vice versa, could be  
20 detected by a change in polarisation. In case that both binding partners, which can be either the full length proteins as well as one binding partner as the full length protein and the other just represented as a peptide are fluorescently labeled, binding could be detected by fluorescence energy transfer (FRET) from one fluorophore to the other. In addition, a variety of commercially  
25 available assay principles suitable for detection of protein-protein interaction are well known in the art, for example but not exclusively AlphaScreen (PerkinElmer) or Scintillation Proximity Assays (SPA) by Amersham. Alternatively, the interaction of the proteins of the invention with cellular proteins could be the basis for a cell-based screening assay, in which both  
30 proteins are fluorescently labeled and interaction of both proteins is detected by analysing cotranslocation of both proteins with a cellular imaging reader, as has been developed for example, but not exclusively, by Cellomics or EvotecOAI. In all cases the two or more binding partners can be different

- 38 -

proteins with one being the protein of the invention, or in case of dimerization and/or oligomerization the protein of the invention itself. Proteins of the invention, for which one target mechanism of interest, but not the only one, would be such protein-protein interaction are *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous proteins.

Assays for determining enzymatic activity of the proteins of the invention are well known in the art. Well known in the art are also a variety of assay formats to measure receptor-ligand binding or receptor downstream signalling.

Genetic reporter systems are widely used to study eukaryotic gene expression and cellular physiology. Applications include the study of receptor activity, transcription factors, intracellular signalling, mRNA processing, and protein folding. For example, the firefly luciferase is used as a reporter because the luciferase assay is very sensitive and rapid. Luciferase reporter assays are commercially available, e.g. from BD Bioscience, Promega, and Boehringer Mannheim. Other reporter genes can be also used to detect eukaryotic gene expression, like chloramphenicol acetyltransferase (CAT), beta-galactosidase (beta-Gal), or human placental alkaline phosphatase (SEAP).

Of particular interest are screening assays for agents that have a low toxicity for mammalian cells. The term "agent" as used herein describes any molecule, e.g. protein or pharmaceutical, with the capability of altering or mimicking the physiological function of one or more of the proteins of the invention. Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 Daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise carbocyclic or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above

- 39 -

functional groups.

Candidate agents are also found among biomolecules including peptides, saccharides, fatty acids, steroids, purines, pyrimidines, nucleic acids and derivatives, structural analogs or combinations thereof. Candidate agents are  
5 obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides and  
10 oligopeptides. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known  
15 pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs. Where the screening assay is a binding assay, one or more of the molecules may be joined to a label, where the label can directly or indirectly provide a detectable signal.

20

Candidate agents may also be found in kinase or phosphatase assays where a kinase or phosphatase substrate such as a protein, a peptide, a lipid, or an organic compound, which may or may not include modifications as further described below, or others are (de)phosphorylated by the proteins or protein  
25 fragments of the invention. The kinase or phosphatase can be a protein of the invention (e.g. *Cdk4*, *fwd*, or *Adk3* homologous kinase or CG7134 or *Pp2C1* homologous phosphatase) or a kinase or phosphatase which is influenced in its activity by a protein of the invention. A therapeutic candidate agent may be identified by its ability to increase or decrease the enzymatic activity of the  
30 proteins of the invention. The kinase or phosphatase activity may be detected by change of the chemical, physical or immunological properties of the substrate due to (de)phosphorylation. One example could be the transfer of radioisotopically labelled phosphate groups from an appropriate donor

- 40 -

molecule to the kinase substrate, or the cleavage of radioisotopically labelled phosphate groups from a phosphatase substrate, catalyzed by the polypeptides of the invention. The (de)phosphorylation of the substrate may be followed by detection of the substrates autoradiography with techniques well known in the art.

Yet in another example, the change of mass of the substrate due to its (de) phosphorylation may be detected by mass spectrometry techniques. One could also detect the phosphorylation status of a substrate with an analyte discriminating between the phosphorylated and unphosphorylated status of the substrate. Such an analyte may act by having different affinities for the phosphorylated and unphosphorylated forms of the substrate or by having specific affinity for phosphate groups. Such an analyte could be, but is not limited to, an antibody or antibody derivative, a recombinant antibody-like structure, a protein, a nucleic acid, a molecule containing a complexed metal ion, an anion exchange chromatography matrix, an affinity chromatography matrix or any other molecule with phosphorylation dependend selectivity towards the substrate.

Such an analyte could be employed to detect the kinase or phosphatase substrate, which is immobilized on a solid support during or after an enzymatic reaction. If the analyte is an antibody, its binding to the substrate could be detected by a variety of techniques as they are described in Harlow and Lane, 1998, Antibodies, CSH Lab Press, NY. If the analyte molecule is not an antibody, it may be detected by virtue of its chemical, physical or immunological properties, being endogenously associated with it or engineered to it.

Yet in another example the kinase or phosphatase substrate may have features, designed or endogenous, to facilitate its binding or detection in order to generate a signal that is suitable for the analysis of the substrates phosphorylation status. These features may be, but are not limited to, a biotin molecule or derivative thereof, a glutathione-S-transferase moiety, a moiety of



- 41 -

six or more consecutive histidine residues, an amino acid sequence or hapten to function as an epitope tag, a fluorochrome, an enzyme or enzyme fragment. The kinase or phosphatase substrate may be linked to these or other features with a molecular spacer arm to avoid steric hindrance.

5

In one example, the kinase or phosphatase substrate may be labelled with a fluorochrome. The binding of the analyte to the labelled substrate in solution may be followed by the technique of fluorescence polarization as it is described in the literature (see, for example, Deshpande S. et al., (1999) Prog. Biomed. Optics (SPIE) 3603: 261; Parker, G.J. et al. (2000) J. Biomol. Screen. 5: 77-88; Wu P. et al., (1997) Anal. Biochem. 249: 29-36). In a variation of this example, a fluorescent tracer molecule may compete with the substrate for the analyte to detect kinase or phosphatase activity by a technique which is known to those skilled in the art as indirect fluorescence polarization.

15

In a further example, candidate agents might also be found in kinase assays that utilize bioluminescent measurement of ATP (*fwd* and *Adk3* homologous proteins of the invention). The bioluminescent kinase assays are commercially available, e.g. from Bio-Whittaker.

20

Another technique for drug screening, which may be used, provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to the proteins of the invention large numbers of different small test compounds, e.g. aptamers, peptides, low-molecular weight compounds etc., are provided or synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with a protein of the invention or fragments thereof, and washed. Bound proteins are then detected by methods well known in the art. Purified proteins can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

25

30

- 42 -

In another embodiment, one may use competitive drug screening assays in which neutralising antibodies capable of binding the protein specifically compete with a test compound for binding the *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous proteins. In this manner, the antibodies  
5 can be used to detect the presence of any peptide, which shares one or more antigenic determinants with the *fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous proteins.

10 The nucleic acids encoding the protein of the invention can be used to generate transgenic animals or site-specific gene modifications in cell lines. These transgenic non-human animals are useful in the study of the function and regulation of the protein of the invention in vivo. Transgenic animals, particularly mammalian transgenic animals, can serve as a model system for  
15 the investigation of many developmental and cellular processes common to humans. A variety of non-human models of metabolic disorders can be used to test effectors/modulators of the protein of the invention. Misexpression (for example, overexpression or lack of expression) of the protein of the invention, particular feeding conditions, and/or administration of biologically  
20 active compounds can create models of metabolic disorders.

In one embodiment of the invention, such assays use mouse models of insulin resistance and/or diabetes, such as mice carrying gene knockouts in the leptin pathway (for example, *ob* (leptin) or *db* (leptin receptor) mice).  
25 Such mice develop typical symptoms of diabetes, show hepatic lipid accumulation and frequently have increased plasma lipid levels (see Bruning J.C. et al., 1998, supra). Susceptible wild type mice (for example C57Bl/6) show similar symptoms if fed a high fat diet. In addition to testing the expression of the proteins of the invention in such mouse strains (see  
30 Examples section), these mice could be used to test whether administration of a candidate effector/modulator alters for example lipid accumulation in the liver, in plasma, or adipose tissues using standard assays well known in the art, such as FPLC, colorimetric assays, blood glucose level tests, insulin

- 43 -

tolerance tests and others.

Transgenic animals may be made through homologous recombination in non-human embryonic stem cells, where the normal locus of the gene encoding the protein of the invention is altered. Alternatively, a nucleic acid construct encoding the protein of the invention is injected into oocytes and is randomly integrated into the genome. Vectors for stable integration include plasmids, retroviruses and other animal viruses, yeast artificial chromosomes (YACs), and the like. The modified cells or animals are useful in the study of the function and regulation of the protein of the invention. For example, a series of small deletions and/or substitutions may be made in the gene that encodes the protein of the invention to determine the role of particular domains of the protein, functions in pancreatic differentiation, etc.

Furthermore, variants of the gene of the invention like specific constructs of interest include anti-sense molecules, which will block the expression of the protein of the invention, or expression of dominant negative mutations. A detectable marker, such as for example lac-Z or luciferase may be introduced in the locus of the gene of the invention, where up regulation of expression of the gene of the invention will result in an easily detected change in phenotype.

One may also provide for expression of the gene of the invention or variants thereof in cells or tissues where it is not normally expressed or at abnormal times of development. In addition, by providing expression of the protein of the invention in cells in which they are not normally produced, one can induce changes in cell behavior.

DNA constructs for homologous recombination will comprise at least portions of the gene of the invention with the desired genetic modification, and will include regions of homology to the target locus. DNA constructs for random integration do not need to contain regions of homology to mediate recombination. Conveniently, markers for positive and negative selection are included. DNA constructs for random integration will consist of the nucleic

- 44 -

acids encoding the protein of the invention, a regulatory element (promoter), an intron and a poly-adenylation signal. Methods for generating cells having targeted gene modifications through homologous recombination are known in the art. For non-human embryonic stem (ES) cells, an ES cell line may be employed, or embryonic cells may be obtained freshly from a host, e.g. mouse, rat, guinea pig, etc. Such cells are grown on an appropriate fibroblast-feeder layer and are grown in the presence of leukemia inhibiting factor (LIF).

When non-human ES or embryonic cells or somatic pluripotent stem cells have been transfected, they may be used to produce transgenic animals. After transfection, the cells are plated onto a feeder layer in an appropriate medium. Cells containing the construct may be selected by employing a selective medium. After sufficient time for colonies to grow, they are picked and analyzed for the occurrence of homologous recombination or integration of the construct. Those colonies that are positive may then be used for embryo transfection and morula aggregation. Briefly, morulae are obtained from 4 to 6 week old superovulated females, the Zona Pellucida is removed and the morulae are put into small depressions of a tissue culture dish. The ES cells are trypsinized, and the modified cells are placed into the depression closely to the morulae. On the following day the aggregates are transferred into the uterine horns of pseudopregnant females. Females are then allowed to go to term. Chimeric offsprings can be readily detected by a change in coat color and are subsequently screened for the transmission of the mutation into the next generation (F1-generation). Offspring of the F1-generation are screened for the presence of the modified gene and males and females having the modification are mated to produce homozygous progeny. If the gene alterations cause lethality at some point in development, tissues or organs can be maintained as allogenic or congenic grafts or transplants, or in vitro culture. The transgenic animals may be any non-human mammal, such as laboratory animal, domestic animals, etc., for example, mouse, rat, guinea pig, sheep, cow, pig, and others. The transgenic animals may be used in functional studies, drug screening, and

- 45 -

other applications and are useful in the study of the function and regulation of the protein of the invention in vivo.

Finally, the invention also relates to a kit comprising at least one of

- 5 (a) a nucleic acid molecule coding for the protein of the invention or a functional fragment thereof;
- (b) the protein of the invention or a functional fragment or an isoform thereof;
- (c) a vector comprising the nucleic acid of (a);
- 10 (d) a host cell comprising the nucleic acid of (a) or the vector of (c);
- (e) a polypeptide encoded by the nucleic acid of (a);
- (f) a fusion polypeptide encoded by the nucleic acid of (a);
- (g) an antibody, an aptamer or another effector/modulator of the nucleic acid of (a) or the polypeptide of (b), (e), or (f) and
- 15 (h) an anti-sense oligonucleotide of the nucleic acid of (a).

The kit may be used for diagnostic or therapeutic purposes or for screening applications as described above. The kit may further contain user instructions.

20 The Figures show:

**Figure 1** shows the content of energy storage triglyceride (TG) of a *Drosophila fwd* (GadFly Accession Number CG7004) mutant. Shown is the change of triglyceride content of HD-EP(3)30148 flies caused by integration of the  
25 P-vector into the annotated transcription unit ('HD-EP30148 (90°C)', column 3) in comparison to controls containing about 2000 fly lines of the proprietary EP collection ('HD-control (90°C)', column 1) and wildtype controls determined in more than 80 independent assays (referred to as 'WT-control (90°C)', column 2).

30

**Figure 2** shows the molecular organization of the mutated *fwd* gene locus.

- 46 -

**Figure 3** shows the expression of the *fwd* homolog in mammalian (mouse) tissues.

**Figure 3A** shows the real-time PCR analysis of the catalytic beta polypeptide of phosphatidylinositol 4 kinase (Pik4cb) expression in wild-type mouse tissues.

**Figure 3B** shows the real-time PCR analysis of Pik4cb expression in different mouse models.

**Figure 3C** shows the real-time PCR analysis of Pik4cb expression in mice fed with a high fat diet compared to mice fed with a control diet.

**Figure 3D** shows the real-time PCR analysis of Pik4cb expression during the differentiation of 3T3-L1 cells from preadipocytes to mature adipocytes.

**Figure 4** shows the expression of the human *fwd* homolog in mammalian (human) tissue.

**Figure 4A** shows the microarray analysis of phosphatidylinositol 4-kinase, catalytic, beta polypeptide (PIK4CB) expression in human abdominal derived primary adipocyte cells during the differentiation from preadipocytes to mature adipocytes.

**Figure 4B** shows the microarray analysis of PIK4CB expression in a human adipocyte cell line during the differentiation from preadipocytes to mature adipocytes.

**Figure 5** shows the content of energy storage glycogen of a *Drosophila Pp2C1* (GadFly Accession Number CG2984) mutant. Shown is the change of glycogen content of HD-EP(X)10310 flies caused by integration of the P-vector into the annotated transcription unit (columns 2 and 4) in comparison to controls containing about 20 fly lines of the proprietary EP collection ('HD-control (70°C)'), column 1, 'HD-control (90°C)'), column 3).

**Figure 6** shows the molecular organization of the mutated *Pp2C1* gene locus.

**Figure 7** shows the expression of the *Pp2C1* homolog in mammalian (mouse) tissues.

- 47 -

**Figure 7A** shows the real-time PCR analysis of the magnesium-dependent protein phosphatase delta isoform (Ppm1d) expression in wild-type mouse tissues.

**Figure 7B** shows the real-time PCR analysis of Ppm1d expression in different mouse models.

**Figure 7C** shows the real-time PCR analysis of Ppm1d expression in mice fed with a high fat diet compared to mice fed with a control diet..

**Figure 7D** shows the real-time PCR analysis of Ppm1d expression during the differentiation of 3T3-L1 cells from preadipocytes to mature adipocytes.

10

**Figure 8** shows the content of energy storage triglyceride (TG) of a *Drosophila Adk3* (GadFly Accession Number CG6612) mutant. Shown is the change of triglyceride content of HD-EP(3)36627 flies caused by integration of the P-vector into the annotated transcription unit ('HD-36627/TM3 (90°C)', column 3) in comparison to controls containing about 2000 fly lines of the proprietary EP collection ('HD-control (90°C)', column 1) and wildtype controls determined in more than 80 independent assays (referred to as 'WT-control (90°C)', column 2). Also shown is the change of triglyceride content of HD-EP(3)36627 flies at different assay conditions (70°C instead of 90°C) ('HD-36627/TM3 (70°C)', column 6) in comparison to controls containing about 880 fly lines of the proprietary EP collection ('HD-control (70°C)', column 4) and wildtype controls determined in 4 independent assays (referred to as 'WT-control (70°C)', column 5).

15

20

**Figure 9** shows the molecular organization of the mutated *Adk3* gene locus.

25

**Figure 10** shows the expression of an *Adk3* homolog in mammalian (mouse) tissues.

30

**Figure 10A** shows the real-time PCR analysis of adenylate kinase 3 alpha like (Akl3l) expression in wild-type and control-diet mouse tissues.

**Figure 10B** shows the real-time PCR analysis of Akl3l expression in different mouse models compared to wild type mice and in mice fed with a high fat diet compared to mice fed with a control diet.

- 48 -

**Figure 10C** shows the real-time PCR analysis of *Akl3l* expression during the differentiation of 3T3-L1 cells from preadipocytes to mature adipocytes.

**Figure 10D** shows the real-time PCR analysis of adenylate kinase 4 (*Ak4*) in wild type and control-diet mouse tissues.

5 **Figure 10E** shows the real-time PCR analysis of *Ak4* expression in different mouse models and in mice fed with a high fat diet compared to mice fed with a control diet

**Figure 10F** shows the real-time PCR analysis of *Ak4* expression in 3T3-L1 cells differentiated from preadipocytes to mature adipocytes.

10

**Figure 11** shows the expression of human *Adk3* homologs in mammalian (human) tissue.

**Figure 11A** shows the microarray analysis of adenylate kinase 3 like 1 (*AK3L1*) expression in human abdominal derived primary adipocyte cells during the differentiation from preadipocytes to mature adipocytes.

15

**Figure 11B** shows the microarray analysis of adenylate kinase 3 (*AK3*) expression in a human adipocyte cell line during the differentiation from preadipocytes to mature adipocytes.

20 **Figure 12** shows the content of energy storage triglyceride (TG) of a *Drosophila* CG3860 (*GadFly* Accession Number) mutant. Shown is the change of triglyceride content of HD-EP(2)25831 flies caused by integration of the P-vector into the annotated transcription unit ('HD-EP25831 (90°C)', column 3) in comparison to controls containing about 2000 fly lines of the proprietary EP collection ('HD-control (90°C)', column 1) and wildtype controls determined in more than 80 independent assays (referred to as 'WT-control (90°C)', column 2).

25

30

**Figure 13** shows the molecular organization of the mutated CG3860 gene locus.

**Figure 14** shows the expression of CG3860 homologs in mammalian (mouse) tissues.



- 49 -

**Figure 14A** shows the real-time PCR analysis of oxysterol binding protein-like 1A (Osbp1a) expression in wild-type and control-diet mouse tissues.

**Figure 14B** shows the real-time PCR analysis of Osbp1a expression in different mouse models compared to wild type mice and in mice fed with a high fat diet compared to mice fed with a control diet..

**Figure 14C** shows the real-time PCR analysis of Osbp1a expression during the differentiation of 3T3-L1 cells from preadipocytes to mature adipocytes.

**Figure 14D** shows the real-time PCR analysis of oxysterol binding protein-like 2 (Osbp12) expression in wild-type and control-diet mouse tissues.

**Figure 14E** shows the real-time PCR analysis of Osbp12 expression in different mouse models compared to wild type mice and in mice fed with a high fat diet compared to mice fed with a control diet..

**Figure 14F** shows the real-time PCR analysis of Osbp12 expression during the differentiation of 3T3-L1 cells from preadipocytes to mature adipocytes.

**Figure 15** shows the expression of human CG3860 homologs in mammalian (human) tissue.

**Figure 15A** shows the microarray analysis of oxysterol binding protein-like 1A (OSBPL1A) expression in human abdominal derived primary adipocyte cells during the differentiation from preadipocytes to mature adipocytes.

**Figure 15B** shows the microarray analysis of OSBPL1A expression in a human adipocyte cell line during the differentiation from preadipocytes to mature adipocytes.

**Figure 15C** shows the microarray analysis of oxysterol binding protein-like 2 (OSBPL2) expression in a human adipocyte cell line during the differentiation from preadipocytes to mature adipocytes.

**Figure 16** shows the content of energy storage triglycerides of a *Drosophila* Cdk4 (GadFly Accession Number CG5072) mutant. Shown is the change of triglyceride content of HD-EP(2)21120 flies caused by integration of the P-vector into the annotated transcription unit (column 2) in comparison to controls containing all fly lines of the proprietary EP collection ('EP-control', column 1)

**Figure 17** shows the molecular organization of the mutated *Cdk4* gene locus.

**Figure 18** shows the content of triglyceride of a *Drosophila* CG7134 (GadFly  
5 Accession Number) mutant. Shown is the change of triglyceride content of  
HD-EP(2)20271 flies caused by integration of the P-vector into the annotated  
transcription unit (column 3) in comparison to controls containing more than  
2000 fly lines of the proprietary EP collection ('HD-control (TG)', column 1) and  
wildtype controls determined in more than 80 independent assays (referred to  
10 as 'WT-control (TG)', column 2).

**Figure 19** shows the molecular organization of the mutated CG7134 gene  
locus.

**Figure 20** shows the content of energy storage triglycerides of a *Drosophila*  
15 *Eip75B* (GadFly Accession Number CG8127) mutant. Shown is the change  
of triglyceride content of HD-EP(3)30293 flies caused by integration of the  
P-vector into the annotated transcription unit (column 2) in comparison to  
controls containing all fly lines of the proprietary EP collection ('EP-control',  
20 column 1)

**Figure 21** shows the molecular organization of the mutated *Eip75B* gene  
locus.

**Figure 22** shows the expression of *Eip75* homologs in different mammalian  
25 models.

**Figure 22A** shows the real-time PCR analysis of nuclear receptor subfamily 1,  
group D, member 1 (Nr1d1) expression in wildtype mouse tissues.

**Figure 22B** shows the real-time PCR analysis of Nr1d1 expression in different  
30 mouse models.

**Figure 22C** shows real-time PCR analysis of Nr1d1 expression during  
differentiation of mammalian fibroblast (3T3-L1) cells from pre-adipocytes to  
mature adipocytes.

- 51 -

**Figure 22D** shows the real-time PCR analysis of nuclear receptor subfamily 1, group D, member 2 (Nr1d2) expression in wildtype mouse tissues.

**Figure 22E** shows the real-time PCR analysis of Nr1d2 expression in different mouse models.

5 **Figure 22F** shows real-time PCR analysis of Nr1d2 expression during differentiation of mammalian fibroblast (3T3-L1) cells from pre-adipocytes to mature adipocytes.

10 **Figure 23** shows the real-time PCR analysis of the expression of *Eip75B* homologs in different human tissues.

**Figure 23A** shows the real-time PCR analysis of NR1D1 expression in different human tissues.

**Figure 23B** shows the real-time PCR analysis of NR1D1 expression in different human adipose tissues.

15 **Figure 23C** shows the real-time PCR analysis of NR1D1 expression in human abdominal derived primary adipocytes during preadipocyte differentiation.

**Figure 23D** shows the real-time PCR analysis of NR1D2 expression in different human tissues.

20 **Figure 23E** shows the real-time PCR analysis of NR1D2 expression in different human adipose tissues.

**Figure 23F** shows the real-time PCR analysis of NR1D2 expression in human abdominal derived primary adipocytes during preadipocyte differentiation.

25

The examples illustrate the invention:

**Example 1: Measurement of energy storage metabolites (ESM; triglyceride and/or glycogen) content in *Drosophila***

30

Mutant flies are obtained from a fly mutation stock collection. The flies are grown under standard conditions known to those skilled in the art. In the course of the experiment, additional feedings with bakers yeast

- 52 -

(*Saccharomyces cerevisiae*) are provided. The average change of triglyceride and/or glycogen (herein referred to as energy storage metabolites, ESM) content of *Drosophila* containing the EP-vectors in homozygous viable, hemizygous viable, or homozygous lethal/heterozygous viable integration, was investigated in comparison to control flies grown under the same conditions (see Figures 1, 5, 8, 12, 16, 18, and 20). For determination of triglyceride and glycogen content, flies were incubated for 5 min at 70°C or 90°C in an aqueous buffer using a waterbath, followed by hot extraction. After another 5 min incubation at 70°C or 90°C and mild centrifugation, the triglyceride content of the flies extract was determined using Sigma Triglyceride (INT 336-10 or -20) assay by measuring changes in the optical density according to the manufacturer's protocol. The glycogen content of the flies extract was determined using the Roche Starch UV-method assay (Cat. No. 0207748) by measuring changes in the optical density according to the manufacturer's protocol. As a reference the protein content of the same extract was measured using BIO-RAD DC Protein Assay according to the manufacturer's protocol. These experiments and assays were repeated several times.

The average triglyceride level ( $\mu\text{g}$  triglyceride/ $\mu\text{g}$  protein) of all flies of the EP collection (referred to as 'EP-control') is shown as 100% in the first columns in Figures 16 and 20. The average triglyceride level ( $\mu\text{g}$  triglyceride/ $\mu\text{g}$  protein) of 2108 flies of the proprietary EP collection (referred to as 'HD-control (90°C)') is shown as 100% in the first columns in Figures 1, 8, 12, and 18. The average triglyceride level ( $\mu\text{g}$  triglyceride/ $\mu\text{g}$  protein) of *Drosophila* wildtype strain Oregon R flies determined in 84 independent assays (referred to as 'WT-control (90°C)') is shown as 102% in the second columns in Figure 1, 8, 12, and 18.

The average triglyceride level ( $\mu\text{g}$  triglyceride/ $\mu\text{g}$  protein) of 883 fly lines of the proprietary EP-collection determined at 70°C (referred to as 'HD-control (70°C)') is shown as 100% in the fourth column in Figure 8. The average triglyceride level of *Drosophila* wildtype strain Oregon R flies determined in 4 independent assays at 70°C (referred to as 'WT-control (70°C)') is shown as

- 53 -

116% in the fifth column in Figure 8. The average glycogen level ( $\mu\text{g}$  glycogen/ $\mu\text{g}$  protein) of 19 fly lines of the proprietary EP-collection with X-chromosomal insertions, determined at 70°C (referred to as 'HD-control (70°C)') is shown as 100% in the first column in Figure 5. The average glycogen level ( $\mu\text{g}$  glycogen/ $\mu\text{g}$  protein) of 20 fly lines of the proprietary EP-collection with X-chromosomal insertions, determined at 90°C (referred to as 'HD-control (90°C)') is shown as 100% in the third column in Figure 5. Standard deviations of the measurements are shown as thin bars.

HD-EP(3)30148 homozygous flies ('HD-EP30148 (90°C)', column 3 in Figure 1), HD-EP(3)36627 heterozygous flies ('HD-36627/TM3 (90°C)', column 3, and 'HD-36627/TM3 (70°C)', column 6 in Figure 8), HD-EP(2)25831 homozygous flies ('HD-EP25831 (90°C)', column 3 in Figure 12), HD-EP(2)21120 homozygous flies ('HD-EP21120' column 2 in Figure 16), HD-EP(2)20271 homozygous flies ('HD-20271 (TG)', column 3 in Figure 18), and HD-EP(3)30293 heterozygous flies ('HD-EP30293 / TM3', column 2 in Figure 20) show constantly a higher triglyceride content than the controls. HD-EP(X)10310 hemizygous flies show constantly a slightly higher triglyceride content than the controls (data not shown). HD-EP(X)10310 hemizygous flies ('HD-10310 (70°C)', column 2 in Figure 5, and 'HD-10310 (90°C)', column 4 in Figure 5) show constantly a higher glycogen content than the controls, determined at different temperatures. Therefore, the loss of gene activity is responsible for changes in the metabolism of the energy storage metabolites.

## **Example 2: Identification of Drosophila genes for changes in triglyceride and/or glycogen levels**

Nucleic acids encoding the proteins of the present invention were identified using a plasmid-rescue technique. Genomic DNA sequences were isolated that are localized adjacent to the EP vectors (herein HD-EP(3)30148, HD-EP(X)10310, HD-EP(3)36627, HD-EP(2)25831, HD-EP(2)21120, HD-EP(2)20271, or HD-EP(3)30293) integration. Using those isolated genomic sequences public databases like Berkeley Drosophila Genome Project

- 54 -

(GadFly) were screened, thereby identifying the integration sites of the vectors, and the corresponding genes. The molecular organization of these gene loci is shown in Figures 2, 6, 9, 13, 17, 19, and 21.

5 In Figures 2, 9, 17, and 21, genomic DNA sequence is represented by the assembly as a horizontal black scaled double-headed arrow that includes the integration sites of the vectors for lines HD-EP(3)30148, HD-EP(3)36627, HD-EP(2)21120, or HD-EP(3)30293. Ticks represent the length in basepairs of the genomic DNA (1000 base pairs (Figures 2, 9, and 17) or 10000 base pairs  
10 (Figure 21) per tick). The part of the figure above the double-headed arrow represents the sense strand, the part below the arrow represents the antisense strand. The grey arrows in the upper part of the figures represent BAC clones, the black arrows in the topmost part of the figures represent the sections of the chromosomes. The insertion sites of the P-elements in the *Drosophila* lines are shown as triangles and are labeled. The cDNA sequences of the predicted  
15 genes (as predicted by the Berkeley *Drosophila* Genome Project, GadFly and by Magpie) are shown as dark grey bars (exons), linked by dark grey lines (introns), and are labeled (see also key at the bottom of the figures).

20 The HD-EP(3)30148 vector is homozygous viable integrated 122 base pairs 5' of transcription variant CG7004-RA of the *Drosophila fwd* gene in antisense orientation. The chromosomal localization site of integration of the vector of HD-EP(3)30148 is at gene locus 3L, 61C1. In Figure 2, the transcript variants of the cDNA of the *Drosophila* gene *fwd* are shown in the lower half of the figure, and are labeled. The integration site of HD-EP(3)30148 is indicated with  
25 a black triangle 5' of the first exon of the predicted transcript variants of the *fwd* gene.

The HD-EP(3)36627 vector is homozygous lethal/heterozygous viable  
30 integrated about 350 base pairs 5' of the *Drosophila Adk3* gene in sense orientation. The chromosomal localization site of integration of the vector of HD-EP(3)36627 is at gene locus 3R, 86C7. In Figure 9, the transcript variants of the cDNA of the *Drosophila Adk3* gene are shown in the lower half of the

- 55 -

figure and are labeled. The integration site of HD-EP(3)36627 is indicated with a black triangle 5' of the first exon of the predicted transcript variants of the *Adk3* gene.

5 The HD-EP(2)21120 vector is homozygous viable integrated into an intron of the cDNA of the *Cdk4* gene in sense orientation. The chromosomal localization site of integration of the vector of HD-EP(2)21120 is at gene locus 2R, 53C9 (according to Flybase and Gadfly release 3). In Figure 17, the transcript variants of the cDNA of the *Drosophila Cdk4* gene are shown in the middle of the figure, and are labeled. The integration site of HD-EP(2) 10 21120 is indicated with a black triangle within an intron of the predicted transcript variants of the *Cdk4* gene.

The HD-EP(3)30293 vector is homozygous lethal/heterozygous viable 15 integrated into the third exon of the transcript variant CG8127-RA of the gene *Eip75B* in antisense orientation. The chromosomal localization site of integration of the vector of HD-EP(3)30293 is at gene locus 3L, 75A8-B2. In Figure 21, the transcript variants of the *Drosophila* gene *Eip75B* are shown in the lower half of the figure and are labeled. The integration site of HD-EP 20 (3)30293 is indicated with a black triangle within the first exon of the predicted transcript variant RA of the *Eip75B* gene.

In Figures 6, 13, and 19, genomic DNA sequence is represented by the assembly as a dotted black line in the middle that includes the integration sites 25 of the vectors for lines HD-EP(X)10310, HD-EP(2)25831, or HD-EP(2)20271. Numbers represent the coordinates of the genomic DNA. The upper parts of the figures represent the sense strand "+", the lower parts represent the antisense strand "-". The insertion sites of the P-elements in the *Drosophila* lines are shown as triangles or boxes in the "P-elements +" lines. Transcribed DNA sequences (ESTs) are shown as grey bars in the "EST +", "EST -", "IPI 30 +", and/or the "IPI -" lines, and predicted cDNAs are shown as bars in the "cDNA +" and/ or "cDNA -" lines. Predicted exons of the cDNAs are shown as dark grey bars and predicted introns are shown as light grey bars.

- 56 -

The HD-EP(X)10310 vector is hemizygous viable integrated into an intron of *Pp2C1* in antisense orientation. The chromosomal localization site of integration of the vector of HD-EP(X)10310 is at gene locus X, 4D1 (according to Flybase) or X, 4C8 (according to Gadfly release 2). In Figure 6, numbers represent the coordinates of the genomic DNA (starting at position 4489000 on chromosome X, ending at position 4497000). The insertion site of the P-element in *Drosophila* line HD-EP(X)10310 is labeled. The gene *Pp2C1* (Gadfly Accession Number CG2984) shown in the "cDNA-" line is labeled. The corresponding ESTs are shown "EST -" and the "IPI -" lines.

The HD-EP(2)25831 vector is homozygous viable integrated into base pair 3 of transcription variant CG3860-RA of the CG3860 gene in antisense orientation. The chromosomal localization site of integration of the vector of HD-EP(2)25831 is at gene locus 2R, 60A16-B1. In Figure 13, numbers represent the coordinates of the genomic DNA (starting at position 19035500 on chromosome 2R, ending at position 19039000). The insertion site of the P-element in *Drosophila* line HD-EP(2)25831 is labeled. The gene CG3860 (Gadfly Accession Number) shown in the "cDNA-" line is labeled. The corresponding ESTs are shown "EST -" line.

The HD-EP(2)20271 vector is homozygous viable integrated into the first intron, 823 base pairs 3' of the first exon of the cDNA of the gene CG7134 (GadFly Accession Number) in antisense orientation. The chromosomal localization site of integration of the vector of HD-EP(2)20271 is at gene locus 2L, 28C1 (according to Flybase), or 28C4-5 (according to Gadfly release 3). In Figure 19, numbers represent coordinates of the genomic DNA (starting at position 7793500 on chromosome 2L, ending at position 7804500). The insertion site of the P-element in *Drosophila* line HD-EP(2)20271 is labeled. The gene CG7134 shown in the "cDNA-" line is labeled. The corresponding ESTs are shown "EST -" line is labeled.

Therefore, expression of the cDNAs encoding the proteins of the present



- 57 -

invention could be affected by integration of the vectors, leading to a change in the amount of energy storage triglycerides and/or glycogen.

### Example 3: Identification of human homologous genes and proteins

5

The *Drosophila* genes and proteins encoded thereby with functions in the regulation of triglyceride and/or glycogen metabolism were further analysed using the BLAST algorithm searching in publicly available sequence databases and mammalian homologs were identified (see Table 1).

10

The term "polynucleotide comprising the nucleotide sequence as shown in GenBank Accession number" relates to the expressible gene of the nucleotide sequences deposited under the corresponding GenBank Accession number. The term "GenBank Accession number" relates to NCBI GenBank database entries (Ref.: Benson et al., (2000) *Nucleic Acids Res.* 28: 15-18). The term "IPI Accession Number" relates to ENSEMBL International Protein Index entries (<http://www.ensembl.org/IPI/>; Hubbard T. et al., (2002) *Nucleic Acids Research* 30: 38-41). Sequences homologous to *Drosophila fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* were identified using the publicly available program BLASTP 2.2.3 of the non-redundant protein data base of the National Center for Biotechnology Information (NCBI) (see, Altschul S.F. et al., (1997) *Nucleic Acids Res.* 25: 3389-3402).

15

20

25

30

**Table 1: Human homologs of the *Drosophila* (Dm) genes**

Dm gene Acc. No. Name	Homo sapiens homologous genes and proteins		Name
	Accession Number		
	cDNA	Protein	
CG7004 <i>fwd</i>	NM_002651	NP_002642	phosphatidylinositol 4-kinase, catalytic, beta polypeptide (PIK4CB)
CG2984 <i>Pp2C1</i>	NM_003620	NP_003611	protein phosphatase 1D, magnesium-dependent, delta isoform (PPM1D)
CG6612 <i>Adk3</i>	NM_016282	NP_057366	adenylate kinase 3 like 1
	NM_013410	NP_037542	adenylate kinase 3 (AK3)
CG3860	NM_018030	NP_060500	oxysterol binding protein-like 1A (OSBPL1A), transcript variant OSBPL1A; isoform A
	NM_080597	NP_542164	oxysterol-binding protein-like 1A (OSBPL1A), transcript variant OSBPL1B; isoform B
	NM_133268	NP_579802	oxysterol-binding protein-like 1A (OSBPL1A), transcript variant OSBPL1C; isoform C
	NM_014835	NP_055650	oxysterol binding protein-like protein 2 (OSBPL2), transcript variant 1; isoform 1
	NM_144498	NP_653081	oxysterol-binding protein-like protein 2 (OSBPL2), transcript variant 2; isoform 2
CG5072 <i>Cdk4</i>	NM_000075	NP_000066	cyclin-dependent kinase 4 (CDK4), transcript variant 1; isoform 1
	NM_052984	NP_443710	cyclin-dependent kinase 4 (CDK4), transcript variant 2; isoform 2
	NM_001259	NP_001250	cyclin-dependent kinase 6 (CDK6)
CG7134	NM_003672	NP_003663	CDC14 cell division cycle 14 homolog A ( <i>S. cerevisiae</i> ) (CDC14A), transcript variant 1; isoform 1
	NM_033312	NP_201569	CDC14 cell division cycle 14 homolog A ( <i>S. cerevisiae</i> ) (CDC14A), transcript variant 2; isoform 2
	NM_033313	NP_201570	CDC14 cell division cycle 14 homolog A ( <i>S. cerevisiae</i> ) (CDC14A), transcript variant 3; isoform 3
	NM_003671	NP_003662	CDC14 cell division cycle 14 homolog B ( <i>S. cerevisiae</i> ) (CDC14B), transcript variant 1; isoform 1
	NM_033331	NP_201588	CDC14 cell division cycle 14 homolog B ( <i>S. cerevisiae</i> ) (CDC14B), transcript variant 2; isoform 2
	NM_033332	NP_201589	CDC14 cell division cycle 14 homolog B ( <i>S. cerevisiae</i> ) (CDC14B), transcript variant 3; isoform 3
CG8127 <i>Eip75B</i>	NM_021724	NP_068370	nuclear receptor subfamily 1, group D, member 1 (NR1D1)
	NM_005126	NP_005117	nuclear receptor subfamily 1, group D, member 2 (NR1D2)

- 59 -

*fwd*, *Pp2C1*, *Adk3*, CG3860, *Cdk4*, CG7134, or *Eip75B* homologous proteins and nucleic acid molecules coding therefore are obtainable from insect or vertebrate species, e.g. mammals or birds. Particularly preferred are nucleic acids as described in Table 1.

5

The mouse homologous cDNAs encoding the polypeptides of the invention were identified as GenBank Accession Numbers XM\_196305 and XM\_205921 (for the mouse homologs of PIK4CB), GenBank Accession Number NM\_016910 (for the mouse homolog of PPM1D), XM\_129200 (for the mouse homolog of AK3L1), AB020239 and D85036 (for the mouse homolog of AK3), NM\_020573 (for the mouse homolog of OSBPL1A), NM\_144500 (for the mouse homolog of OSBPL2), NM\_009870 (for the mouse homolog of CDK4), NM\_009873 (for the mouse homolog of CDK6), IPI Accession Numbers IPI00134094 and IPI00111965 (for the mouse homologs of CDC14A), IPI Accession Numbers IPI00126617, IPI00126961, and IPI00124775 (for the mouse homologs of CDC14B), NM\_145434 and XM\_126627 (for the mouse homologs of NR1D1), and NM\_011584 (for the mouse homologs of NR1D2).

#### 20 **Example 4: dUCPy modifier screen**

Expression of *Drosophila* uncoupling protein dUCPy in a non-vital organ like the eye (Gal4 under control of the eye-specific promoter of the "eyeless" gene) results in flies with visibly damaged eyes. This easily visible eye phenotype is the basis of a genetic screen for gene products that can modify UCP activity.

Parts of the genomes of the strain with Gal4 expression in the eye and the strain carrying the pUAST-dUCPy construct were combined on one chromosome using genomic recombination. The resulting fly strain has eyes that are permanently damaged by dUCPy expression. Flies of this strain were crossed with flies of a large collection of mutagenized fly strains. In this mutant collection a special expression system (EP-element, Ref.: Rorth P., (1996) Proc Natl Acad Sci US 93: 12418-12422) is integrated randomly in

- 60 -

different genomic loci. The yeast transcription factor Gal4 can bind to the EP-element and activate the transcription of endogenous genes close the integration site of the EP-element. The activation of the genes therefore occurs in the same cells (eye) that overexpress dUCPy. Since the mutant collection contains several thousand strains with different integration sites of the EP-element it is possible to test a large number of genes whether their expression interacts with dUCPy activity. In case a gene acts as an enhancer of UCP activity the eye defect will be worsened; a suppressor will ameliorate the defect.

Using this screen a gene with enhancing activity was discovered that was found to be the *Eip75B* gene in *Drosophila*.

#### **Example 5: Expression of the polypeptides in mammalian tissues**

##### **Example 5A: Expression of the polypeptides in mouse tissues**

To analyse the expression of the polypeptides disclosed in this invention in mammalian tissues, several mouse strains (preferably mice strains C57Bl/6J, C57Bl/6 ob/ob and C57Bl/KS db/db which are standard model systems in obesity and diabetes research) were purchased from Harlan Winkelmann (33178 Borcheln, Germany) and maintained under constant temperature (preferably 22°C), 40 per cent humidity and a light / dark cycle of preferably 14 / 10 hours. The mice were fed a standard chow (for example, from ssniff Spezialitäten GmbH, order number ssniff M-Z V1126-000). For the fasting experiment ("fasted wild type mice"), wild type mice were starved for 48 h without food, but only water supplied ad libitum (see, for example, Schnetzler B. et al., (1993) J Clin Invest 92: 272-280, Mizuno T.M. et al., (1996) Proc Natl Acad Sci USA 93: 3434-3438). In a further experiment wild-type (wt) mice were fed a control diet (preferably Altromin C1057 mod control, 4.5% crude fat) or high fat diet (preferably Altromin C1057mod. high fat, 23.5% crude fat). Animals were sacrificed at an age of

- 61 -

6 to 8 weeks. The animal tissues were isolated according to standard procedures known to those skilled in the art, snap frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until needed.

5 For analyzing the role of the proteins disclosed in this invention in the *in vitro* differentiation of mammalian cell culture cells for the conversion of pre-adipocytes to adipocytes, mammalian fibroblast (3T3-L1) cells (e.g., Green H. and Kehinde O., (1974) Cell 1: 113-116) were obtained from the American Tissue Culture Collection (ATCC, Hanassas, VA, USA; ATCC- CL  
10 173). 3T3-L1 cells were maintained as fibroblasts and differentiated into adipocytes as described in the prior art (e.g., Qiu Z. et al., (2001) J. Biol. Chem. 276: 11988-11995; Sliker L.J. et al., (1998) BBRC 251: 225-229). In brief, cells were plated in DMEM/10% FCS (Invitrogen, Karlsruhe, Germany) at 50,000 cells/well in duplicates in 6-well plastic dishes and cultured in a  
15 humidified atmosphere of 5%  $\text{CO}_2$  at  $37^{\circ}\text{C}$ . At confluence (defined as day 0: d0) cells were transferred to serum-free (SF) medium, containing DMEM/HamF12 (3:1; Invitrogen), fetuin (300  $\mu\text{g}/\text{ml}$ ; Sigma, Munich, Germany), transferrin (2  $\mu\text{g}/\text{ml}$ ; Sigma), pantothenate (17 $\mu\text{M}$ ; Sigma), biotin (1 $\mu\text{M}$ ; Sigma), and EGF (0.8nM; Hoffmann-La Roche, Basel, Switzerland).  
20 Differentiation was induced by adding dexamethasone (DEX; 1 $\mu\text{M}$ ; Sigma), 3-methyl-isobutyl-1-methylxanthine (MIX; 0.5mM; Sigma), and bovine insulin (5 $\mu\text{g}/\text{ml}$ ; Invitrogen). Four days after confluence (d4), cells were kept in SF medium, containing bovine insulin (5 $\mu\text{g}/\text{ml}$ ) until differentiation was completed. At various time points of the differentiation procedure, beginning  
25 with day 0 (day of confluence) and day 2 (hormone addition; for example, dexamethasone and 3-isobutyl-1-methylxanthine), up to 12 days of differentiation, suitable aliquots of cells were taken every two days.

RNA was isolated from mouse tissues or cell culture cells using Trizol  
30 Reagent (for example, from Invitrogen, Karlsruhe, Germany) and further purified with the RNeasy Kit (for example, from Qiagen, Germany) in combination with a DNase-treatment according to the instructions of the manufacturers and as known to those skilled in the art. Total RNA was

- 62 -

reverse transcribed (preferably using Superscript II RNaseH<sup>-</sup> Reverse Transcriptase, from Invitrogen, Karlsruhe, Germany) and subjected to Taqman analysis preferably using the Taqman 2xPCR Master Mix (from Applied Biosystems, Weiterstadt, Germany; the Mix contains according to  
5 the Manufacturer for example AmpliTaq Gold DNA Polymerase, AmpErase UNG, dNTPs with dUTP, passive reference Rox and optimized buffer components) on a GeneAmp 5700 Sequence Detection System (from Applied Biosystems, Weiterstadt, Germany).

10 Taqman analysis was performed preferably using the following primer/probe pairs:

For the amplification of mouse phosphatidylinositol 4-kinase, catalytic, beta polypeptide (Pik4cb) sequence (GenBank Accession Number XM\_196305):

15 Mouse Pik4cb forward primer (SEQ ID NO: 1): 5'- TGG AAG CAA GGA AGC TGA GAA C -3'; mouse Pik4cb reverse primer (SEQ ID NO: 2): 5'- GCT CGG AAG CAC ATG GAC A -3'; Pik4cb Taqman probe (SEQ ID NO: 3): (5/6-FAM)- TGC TGC ATG ATC TCC ACA ATC TGT ACC A -(5/6-TAMRA).

20 For the amplification of mouse protein phosphatase 1D magnesium-dependent, delta isoform (Ppm1d) sequence (GenBank Accession Number NM\_016910):

Mouse Ppm1d forward primer (SEQ ID NO: 4): 5'- ATG GCC AAA GAC TAT GAC AGG TC -3'; mouse Ppm1d reverse primer (SEQ ID NO: 5): 5'- TGC  
25 CCC GTA TTA TAA CCA CAC TG -3'; Ppm1d Taqman probe (SEQ ID NO: 6): (5/6-FAM)- TCC CAG CAC GTC CGG GAC AAC T -(5/6-TAMRA).

For the amplification of mouse adenylate kinase 3 alpha like (Akl3l) sequence (GenBank Accession Number XM\_129200):

30 Mouse Akl3l forward primer (SEQ ID NO:7): 5'- CAG TGG TCG AGT ATA CAA CAT TGA ATT -3'; mouse Akl3l reverse primer (SEQ ID NO: 8): 5'- GCT GAA TCA GAG GTT CTC CGG -3'; Akl3l Taqman probe (SEQ ID NO: 9): (5/6-FAM)- CCA AGA CTG TGG GCA TTG ATG ACC TG -(5/6-TAMRA).

For the amplification of mouse adenylate kinase 4 (Ak4) sequence (GenBank

- 63 -

Accession Number D85036):

5 Mouse Ak4 forward primer (SEQ ID NO:10): 5'- CGC CCA GAA CTT TG  
GCC -3'; mouse Ak4 reverse primer (SEQ ID NO: 11): 5'- CGG TCT TGA  
GGT TCT CCC G -3'; Ak4 Taqman probe (SEQ ID NO: 12): (5/6-FAM)- CCA  
GCA TCT CTC CAG CGG CCA CT -(5/6-TAMRA).

For the amplification of mouse oxysterol binding protein-like 1A (Osbp1a)  
sequence (GenBank Accession Number NM\_020573):

10 Mouse Osbp1a forward primer (SEQ ID NO: 13): 5'- TGG CTG CAT TTG  
CTG TGT CT -3'; mouse Osbp1a reverse primer (SEQ ID NO: 14): 5'- CGG  
GTT GAA GGG CTT TCC -3'; Osbp1a Taqman probe (SEQ ID NO: 15):  
(5/6-FAM)- CGC CTC TCA GTG GGA GCG CAC -(5/6-TAMRA).

15 For the amplification of mouse oxysterol binding protein-like 2 (Osbp12)  
sequence (GenBank Accession Number NM\_144500):

20 Mouse Osbp12 forward primer (SEQ ID NO: 16): 5'- CAT GTG TAC CTC ATT  
CAC AAA GCC -3'; mouse Osbp12 reverse primer (SEQ ID NO: 17): 5'-  
GGC AGC GAC AGA CTG CAT T -3'; Osbp12 Taqman probe (SEQ ID NO:  
18): (5/6-FAM)- CAA GTC AGT CCC AGC CGC TGG AG -(5/6-TAMRA).

25 For the amplification of mouse nuclear receptor subfamily 1, group D,  
member 1 (Nr1d1) sequence (GenBank Accession Number NM\_145434):

Mouse Nr1d1 forward primer (SEQ ID NO: 19): 5'- CGG CTC AGC GTC ATA  
ATG AAG -3'; mouse Nr1d1 reverse primer (SEQ ID NO: 20): 5'- AGG CCA  
GGT AGG CGG GTA -3'; Nr1d1 Taqman probe (SEQ ID NO: 21): (5/6-  
FAM)- CTG AAT GGT CTA CGC CAG GGC CC -(5/6-TAMRA).

For the amplification of mouse nuclear receptor subfamily 1, group D,  
member 2 (Nr1d2) sequence (GenBank Accession Number NM\_011584):

30 Mouse Nr1d2 forward primer (SEQ ID NO: 22): 5'- ACT GTG ATG CCA ACG  
GCA -3'; mouse Nr1d2 reverse primer (SEQ ID NO: 23): 5'- CTG TGC GGT  
CAC TCT TCA GAA C -3'; Nr1d2 Taqman probe (SEQ ID NO: 24): (5/6-  
FAM)- TCC CAA GAA CGC TGA TAT CTC TAG CAT CGA -(5/6-TAMRA).

- 64 -

In the figures the relative RNA-expression is shown on the Y-axis. In Figures 3A-C, 7A-C, 10A-B, 10D-E, 14A-B, 14D-E, 22A-B, and 22D-E, the tissues tested are given on the X-axis. "WAT" refers to white adipose tissue, "BAT" refers to brown adipose tissue.

5

In Figures 10A-B, 10D-E, 14A-B, and 14D-E, the panel of the wild type mice tissues comprises WAT, BAT, muscle, liver, pancreas, hypothalamus, brain, testis, colon, small intestine, heart, lung, spleen, and kidney, and the panel of the control diet-mice tissues comprises WAT, BAT, muscle, liver, brain, testis, colon, small intestine, heart, lung, spleen, and kidney. In Figures 3D, 7D, 10C, 10F, 14C, 14F, 22C, and 22F, the X-axis represents the time axis. "d0" refers to day 0 (start of the experiment), "d2" - "d12" refers to day 2 - day 12 of adipocyte differentiation.

10

The function of the proteins of the invention in metabolism was further validated by analyzing the expression of the transcripts in different tissues and by analyzing the role in adipocyte differentiation.

15

Mouse models of insulin resistance and/or diabetes were used, such as mice carrying gene knockouts in the leptin pathway (for example, *ob/ob* (leptin) or *db/db* (leptin receptor/ligand) mice) to study the expression of the proteins of the invention. Such mice develop typical symptoms of diabetes, show hepatic lipid accumulation and frequently have increased plasma lipid levels (see Bruning J.C. et al, (1998) Mol. Cell. 2: 559-569).

20

25

Further expression of the mRNAs encoding the proteins of the invention was also examined in susceptible wild type mice (for example, C57Bl/6) that show symptoms of diabetes, lipid accumulation, and high plasma lipid levels, if fed a high fat diet.

30

Expression profiling studies confirm the particular relevance of the proteins of the present invention as regulators of energy metabolism in mammals.



- 65 -

Taqman analysis revealed that the catalytic beta polypeptide of phosphatidylinositol 4-kinase (Pik4cb) is expressed in several mammalian tissues, showing highest level of expression in hypothalamus and higher levels in further tissues, e.g. WAT, BAT, muscle, liver, brain, testis, colon, small intestine, heart, lung, spleen, and kidney. Furthermore Pik4cb is expressed on a lower but still robust levels in the pancreas of wild type mice as depicted in Figure 3A. We found, for example, that the expression of Pik4cb is up-regulated in the brain of ob/ob mice compared to wildtype mice (see Figure 3B). In wild type mice fed a high fat diet, the expression of Pik4cb is up-regulated in WAT and muscle, as depicted in Figure 3C. We further show in this invention (see Figure 3D) that the Pik4cb mRNA is expressed and slightly up-regulated during the differentiation into mature adipocytes. The regulated expression of Pik4cb in different animal models used to study metabolic disorders, together with the regulation during the differentiation from preadipocytes to mature adipocytes, suggest that it plays an essential role in cellular metabolism.

Taqman analysis revealed highest level of expression of magnesium-dependent protein phosphatase, delta isoform (Ppm1d) in the testis of wild type mice, and high levels of expression in WAT, BAT, muscle, hypothalamus, brain, colon, small intestine, heart, lung, spleen, and kidney. Furthermore Ppm1d is expressed on lower but still robust levels in liver and pancreas of wild type mice as depicted in Figure 7A. We found, for example, that the expression of Ppm1d is up-regulated in the brain, pancreas, kidney, muscle, spleen, and hypothalamus of fasted mice compared to wild type mice. We also found that the expression of Ppm1d is up-regulated in the brain, hypothalamus, and kidney of ob/ob mice compared to wild type mice (see Figure 7B). In wild type mice fed a high fat diet, the expression of Ppm1d is up-regulated in BAT and muscle, and down-regulated in the kidney as depicted in Figure 7C. We further show in this invention (see Figure 7D) that the Ppm1d mRNA is expressed and slightly down-regulated during the differentiation into mature adipocytes. The regulated expression of Ppm1d in different animal models used to study metabolic disorders, together with

- 66 -

the regulation during the differentiation from preadipocytes to mature adipocytes, suggests that this gene plays a central role in energy homeostasis.

5 Taqman analysis revealed that adenylate kinase 3 alpha like (Akl3l) is expressed in several mammalian tissues, showing highest level of expression in BAT, liver, and kidney and higher levels in further tissues, e.g. WAT, muscle, hypothalamus, brain, testis, colon, small intestine, heart, lung, and spleen of wild type and control-diet mice. Furthermore Akl3l is  
10 expressed on lower but still robust levels in the pancreas of wild type mice as depicted in Figure 10A. We found, for example, that the expression of Akl3l is down-regulated in the WAT and muscle of fasted mice compared to wild type mice (see Figure 10B). We show in this invention (see Figure 10C) that the Akl3l mRNA is expressed and regulated during the differentiation  
15 into mature adipocytes. Therefore, the Akl3l protein might play an essential role in adipogenesis. The expression of Akl3l in metabolic active tissues and the regulated expression in tissues of animal models used to study metabolic disorders, together with the regulation during the differentiation from preadipocytes to mature adipocytes, suggests that this gene plays a central  
20 role in energy homeostasis.

Taqman analysis revealed that adenylate kinase 4 (Ak4) is expressed in several mammalian tissues, predominantly in the kidney. Furthermore Ak4 is expressed on lower but still high levels in the WAT, BAT, muscle, liver,  
25 hypothalamus, brain, testis, colon, small intestine, heart, and lung of wild type and control-diet mice as depicted in Figure 10D. Ak4 is expressed on lower but detectable levels in muscle, pancreas, small intestine and spleen as shown in Figure 10D. We found, for example, that the expression of Ak4 is down-regulated in the hypothalamus and kidney, and up-regulated in the  
30 pancreas of fasted mice compared to wild type mice. In the ob/ob mice Ak4 is down-regulated in the WAT and kidney and up-regulated in the pancreas compared to wild type mice (see Figure 10E). In wild type mice fed a high fat diet, the expression of Ak4 is up-regulated in muscle, and down-regulated

- 67 -

in the WAT, small intestine, and lung, and spleen, compared to control diet mice, as depicted in Figure 10E. We show in this invention (see Figure 10F) that the Ak4 mRNA is expressed and regulated during the differentiation into mature adipocytes. Therefore, the Ak4 protein might play an essential role in adipogenesis. The down-regulation of Ak4 expression in the WAT of genetically, as well as diet induced obesity models, suggests that this gene plays a central role in energy homeostasis. This result is supported by the transient down-regulation during the differentiation from preadipocytes to mature adipocytes.

Taqman analysis revealed that oxysterol binding protein-like 1A (Osbp1a) is expressed in several mammalian tissues, showing highest level of expression in hypothalamus, brain, and testis, and higher levels in further tissues, e.g. WAT, BAT, muscle, liver, colon, small intestine, heart, lung, and kidney. Furthermore Osbp1a is expressed on a lower but still robust levels in pancreas and spleen of wild type and control-diet mice as depicted in Figure 14A. We found, for example, that the expression of Osbp1a is down regulated in WAT of fasted mice compared to wild type mice (see Figure 14B). In wild type mice fed a high fat diet, the expression of Osbp1a is up-regulated in muscle, compared to control mice diet as depicted in Figure 14B. We show in this invention (see Figure 14C) that the Osbp1a mRNA is expressed and transiently up-regulated during the differentiation into mature adipocytes. Therefore, the Osbp1a protein might play an essential role in adipogenesis. The expression of Osbp1a in metabolic active tissues of wild type mice, as well as the regulation in different animal models used to study metabolic disorders, suggests that this gene plays a central role in energy homeostasis. This result is supported by the regulation during the differentiation from preadipocytes to mature adipocytes.

Taqman analysis revealed that oxysterol binding protein-like 2 (Osbp12) is expressed in several mammalian tissues, showing highest level of expression in WAT, hypothalamus, brain, and colon. Furthermore Osbp12 is expressed on lower but still robust levels in BAT, muscle, liver, pancreas,

- 68 -

testis, small intestine, heart, lung, spleen, and kidney of wild type and control-diet mice as depicted in Figure 14D. We found, for example, that the expression of *Osbpl2* is slightly down-regulated in the small intestine of fasted compared to wild type mice. (see Figure 14E). We show in this invention (see Figure 14F) that the *Osbpl2* mRNA is expressed during the differentiation into mature adipocytes. The expression of *Osbpl2* in metabolic active tissues, together with the expression during the differentiation from preadipocytes to mature adipocytes, suggests that this gene plays a role in energy homeostasis.

As shown in Figure 22A analysis of the expression of *Nr1d1* in mammalian (mouse) tissues revealed that *Nr1d1* is expressed in most mammalian tissues, showing highest level of expression in BAT, muscle, lung and heart, and higher levels in further tissues, e.g. WAT, liver, hypothalamus, brain, testis, colon, small intestine, spleen, and kidney. Furthermore *Nr1d1* is expressed on a lower but still robust levels in pancreas and bone marrow of wild type mice as depicted in Figure 22A. We found, for example, that the expression of *Nr1d1* is down-regulated in the BAT, small intestine, and bone marrow of fasted mice compared to wild type mice. We also found that the expression of *Nr1d1* is down-regulated in small intestine, and bone marrow of *ob/ob* mice compared to wild type mice (see Figure 22B). As shown in Figure 22C, real time PCR (Taqman) analysis of the expression of *Nr1d1* revealed that *Nr1d1* is transiently up-regulated during 3T3-L1 differentiation. The observed 11-fold up-regulation of this transcriptional regulator during adipogenesis suggests a critical role in adipocyte maturation, making it an interesting candidate gene for treating metabolic disorders.

Taqman analysis revealed that *Nr1d2* is expressed in several mammalian tissues, showing highest level of expression in WAT, BAT, and hypothalamus and higher levels in further tissues, e.g. muscle, liver, brain, testis, colon, small intestine, heart, lung, spleen, and kidney. Furthermore *Nr1d2* is expressed on a lower but still robust levels in the pancreas of wild type mice as depicted in Figure 22D. We found, for example, that the

- 69 -

expression of Nr1d2 is down-regulated in WAT of fasted mice compared to wild type mice. We also found that the expression of Nr1d2 is down-regulated in the WAT, BAT, pancreas, hypothalamus, small intestine, lung, and kidney of ob/ob mice compared to wildtype mice (see Figure 22E). We show in this invention (see Figure 22F) that the Nr1d2 mRNA is expressed and up-regulated during the differentiation into mature adipocytes. Therefore, the Nr1d2 protein might play an essential role in adipogenesis. The expression of Nr1d2 in metabolic active tissues of wild type mice, as well as the regulation in different animal models used to study metabolic disorders, suggests that this gene plays a central role in energy homeostasis. This result is supported by the regulation during the differentiation from preadipocytes to mature adipocytes.

#### **Example 5B: *Eip75B* homologous mRNA expression in human tissues**

Human primary adipocytes were differentiated into mature adipocytes as described by (Hauner et al., (1989) J Clin Invest 84: 1663-1670). Briefly, cells were grown in DMEM/Nutrient Mix F12, 1% PenStrep, 17µM biotin, 33µM pantothenate, 10% none heat inactivated fetal calf serum. On day 0 of differentiation, the medium was changed to DMEM/Nutrient Mix F12, 1% Pen/Strep, 17µM biotin, 33µM pantothenate, 0,01mg/ml transferrin, hydrocortisone, 20nM human insulin, 0,2nM T3, 25nM dexamethasone, 250µM IBMX, 3µM rosiglitazone. On day 4 of differentiation, the medium was changed to DMEM/Nutrient Mix F12 1%Pen/Strep, 17µM biotin, 33µM pantothenate, 0,01mg/ml transferrin, 100nM hydrocortisone, 20nM human insulin, 0,2nM T3. At various time points of the differentiation procedure, beginning with day 0 (day of confluence) and day 4 (hormone addition), up to 14 days of differentiation, suitable aliquots of cells were taken. RNA was isolated from human cells using Trizol Reagent (for example, from Invitrogen, Karlsruhe, Germany) and further purified with the RNeasy Kit (for example, from Qiagen, Germany) in combination with an DNase-treatment according to the instructions of the manufacturers and as known to those skilled in the art.

- 70 -

5 RNAs isolated from different human tissues were obtained from Invitrogen Corp., Karlsruhe, Germany, Stratagene Amsterdam, The Netherlands, or BD Biosciences Clontech,. Palo Alto, CA, USA. (i) total RNA from human normal brain (Invitrogen Corp. Order Number D6030-01); (ii) total RNA from human adult skeletal muscle (Stratagene Order Number 735030); (iii) total RNA from human adult lung (Stratagene Order Number 735020); (iv) total RNA from human normal adipose tissue (Invitrogen Corp. Order Number D6005-01); (v) total RNA from human normal pancreas (Invitrogen Corp. Order Number D6101-01); (vi) total RNA from human adult liver (Stratagene Order Number 735018); (vii) total RNA from human adult testis (BD Biosciences Clontech Order Number 64101-1); (viii) total RNA from human adult placenta (Stratagene Order Number 735026). The RNA was treated with DNase according to the instructions of the manufacturers (for example, from Qiagen, Germany) and as known to those skilled in the art.

20 Total RNA was reverse transcribed (preferably using Superscript II RNaseH<sup>-</sup> Reverse Transcriptase, from Invitrogen, Karlsruhe, Germany) and subjected to Taqman analysis preferably using the 'Taqman 2xPCR Master Mix' (from Applied Biosystems, Weiterstadt, Germany). The Taqman 2xPCR Master Mix contains according to the Manufacturer for example AmpliTaq Gold DNA Polymerase, AmpErase UNG, dNTPs with dUTP, passive reference Rox and optimized buffer components) on a GeneAmp 5700 Sequence Detection System (all obtained from Applied Biosystems, Weiterstadt, Germany).

25 Taqman analysis was performed preferably using the following primer/probe pairs:

30 The primer/probe pair for the amplification of human NR1D1 was obtained from Applied Biosystems, Weiterstadt, Germany (Order Number Hs00253876\_m1).

For the amplification of human NR1D1:

- 71 -

Human NR1D1 forward primer (SEQ ID NO: 25): 5'-3';

human NR1D1 reverse primer (SEQ ID NO: 26): 5'-3';

human NR1D1 Taqman probe (SEQ ID NO: 27): (5/6-FAM)-(5/6-TAMRA).

5     For the amplification of human NR1D2:

human NR1D2 forward primer (SEQ ID NO: 28): 5'- GGT CAC TTG CAA  
AAT GAC ACA TTA G -3'; human NR1D2 reverse primer (SEQ ID NO: 29):  
5'- GGT CGC AGC TGT TCC TGG -3'; human NR1D2 Taqman probe (SEQ  
ID NO: 30): (5/6-FAM)- AGA ACA TCA TGA ACA GAC AGC CTT GC CA -  
10     (5/6-TAMRA).

As shown in Figure 23A, real time PCR (Taqman) analysis of the expression  
of NR1D1 in human tissues revealed that NR1D1 is expressed in all tissues  
analysed with highest levels of expression in brain, and high expression  
15     levels in further tissues, e.g. muscle, adipose tissue, liver, and testis and  
placenta. Furthermore NR1D1 is expressed on lower but still robust levels in  
lung and pancreas. As shown in Figure 23B, NR1D1 is expressed in visceral  
as well as in subcutaneous adipose tissue, with a higher expression level in  
the subcutaneous adipose tissue, which is discussed as the fat storage  
20     adipose tissue. As shown in Figure 23C, NR1D1 is expressed during human  
adipocyte differentiation. The high expression of NR1D1 in metabolic active  
tissues (brain, adipose tissue, and muscle) suggests that this gene plays a  
role in energy homeostasis.

25     As shown in Figure 23D, real time PCR (Taqman) analysis of the expression  
of NR1D2 in human tissues revealed that NR1D2 is expressed in all tissues  
analysed with highest levels of expression in brain, muscle, and adipose  
tissue, and high expression levels in further tissues, e.g. lung, liver, testis,  
and placenta. Furthermore NR1D2 is expressed on lower but still robust  
30     levels in the pancreas. As shown in Figure 23E NR1D2 is expressed in  
visceral as well as in subcutaneous Adipose tissue, with a higher expression  
level in the subcutaneous Adipose tissue, which is discussed as the fat  
storage Adipose tissue. As shown in figure 23E , NR1D2 is expressed and

- 72 -

down-regulated during human adipocyte differentiation. The high expression of NR1D2 in the adipose tissue tissues, as well as the expression of NR1D2 during the differentiation from preadipocytes to mature adipocytes, suggests that this gene plays a role in energy homeostasis.

5

**Example 6. Analysis of the differential expression of transcripts of the proteins of the invention in human tissues**

10 RNA preparation from human primary adipose tissues was done as described in Example 5. The target preparation, hybridization, and scanning was performed as described in the manufactures manual (see Affymetrix Technical Manual, 2002, obtained from Affmetrix, Santa Clara, USA).

15 In Figures 4, 11, and 15 the X-axis represents the time axis, shown are day 0 and day 12 of adipocyte differentiation. The Y-axis represents the flourescent intensity. The expression analysis (using Affymetrix GeneChips) of the phosphatidylinositol 4-kinase, catalytic, beta polypeptide (PIK4CB), adenylate kinase 3 like 1 (AK3L1), adenylate kinase 3 (AK3), oxysterol  
20 binding protein-like 1A (OSBPL1A), and oxysterol binding protein-like 2 (OSBPL2) genes using human abdominal derived primary adipocytes and human adipocyte cell line (SGBS) differentiation, clearly shows differential expression of human PIK4CB, AK3L1, AK3, OSBPL1A, and OSBPL2 genes in adipocytes. Several independent experiments were done. The  
25 experiments further show that the PIK4CB, AK3L1, AK3, and OSBPL1A, and transcripts are most abundant at day 12 compared to day 0 during differentiation (see Figures 4, 11, and 15A-B). The experiments further show that the OSBPL2 transcript is most abundant at day 0 compared to day 12 during differentiation (see Figure 15C).

30

Thus, the PIK4CB, AK3L1, AK3, and OSBPL1A proteins have to be significantly increased in order for the preadipocytes to differentiate into mature adipocyte, and the OSBPL2 protein has to be significantly



- 73 -

decreased in order for the preadipocytes to differentiate into mature adipocyte. Therefore, PIK4CB, AK3L1, AK3, and OSBPL1A in preadipocytes have the potential to enhance adipose differentiation, and OSBPL2 in preadipocytes has the potential to inhibit adipose differentiation.

Therefore, PIK4CB, AK3L1, AK3, OSBPL1A, and OSBPL2 proteins might play an essential role in the regulation of human metabolism, in particular in the regulation of adipogenesis and thus it might play an essential role in obesity, diabetes, and/or metabolic syndrome.

#### **Example 7: Generation and analysis of transgenic mice**

Mouse cDNA is isolated from mouse brown adipose tissue (BAT) using standard protocols as known to those skilled in the art. The cDNA is amplified by RT-PCR and point mutations are introduced into the cDNA.

The resulting mutated cDNA is cloned into a suitable transgenic expression vector. The transgene is microinjected into the male pronucleus of fertilized mouse embryos (preferably strain C57/BL6/CBA F1 (Harlan Winkelmann)). Injected embryos are transferred into pseudo-pregnant foster mice. Transgenic founders are detected by PCR analysis. Two independent transgenic mouse lines containing the construct are established and kept on a C57/BL6 background. Briefly, founder animals are backcrossed with C57/BL6 mice to generate F1 mice for analysis. Transgenic mice are continuously bred onto the C57/BL6 background. The expression of the proteins of the invention can be analyzed by taqman analysis as described above, and further analysis of the mice can be done as known to those skilled in the art.

For the purpose of the present invention, it will be understood by the person having average skill in the art that any combination of any feature mentioned

- 74 -

throughout the specification is explicitly disclosed herewith.